Probability density of the wavelet coefficients of a noisy chaos

Matthieu Garcin, Dominique Guegan

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Probability density of the wavelet coefficients of a noisy chaos

Matthieu GARCIN, Dominique GUEGAN

2013.15

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

Probability density of the wavelet coefficients of a noisy chaos

Matthieu Garcin* Dominique Guégan†

January 23, 2013

Abstract

We are interested in the random wavelet coefficients of a noisy signal when this signal is the unidimensional or multidimensional attractor of a chaos. More precisely we give an expression for the probability density of such coefficients. If the noise is a dynamic noise, then our expression is exact. If we face a measurement noise, then we propose two approximations, using Taylor expansion or Edgeworth expansion. We give some illustrations of these theoretical results for the logistic map, the tent map and the Hénon map, perturbed by a Gaussian or a Cauchy noise.

Keywords: wavelets, dynamical systems, chaos, noise, alpha-stable.

1 Introduction

Many methods have been proposed to recover a pure linear signal form noisy observations using wavelets. However, when the signal is non-linear, no satisfying solution has been proposed for the denoising problem. Because reality is often non-linear, we aim to propose a method to denoise those signals, in particular when they are chaotic systems.

We model the attractor of a chaotic system (x_t) by the relation:

$$x_t = z\left(x_{t-1}\right), \tag{1}$$

where z is a real function with a bounded support, denoted $Supp(z)$. Moreover, we suppose that this chaotic system is such that each small interval contained in its support is frequently visited by its trajectory. Thanks to that assumption, interpolation is not mandatory when integrating on a large quantity of observed data generated by the chaotic system.

Such a chaotic system may be perturbed by some noise: in that case, the observable data is not x_t but u_t and the observed attractor is not z but a certain z^{ε} defined by:

$$u_t = z^{\varepsilon}\left(u_{t-1}\right).$$

The observable data, u_t, may be defined by different approaches. Indeed, there exists several ways to model the noise influence. We will consider some kinds of noise, following [12]:

- ▷ Measurement noise is a set of independent identically distributed random variables (ε_t) such that:

$$\begin{cases} u_t &= x_t + \varepsilon_t \\ x_t &= z\left(x_{t-1}\right). \end{cases} \tag{2}$$

*Corresponding author, Université Paris 1 Panthéon-Sorbonne, MSE - CES, 106 boulevard de l'hôpital, 75013 Paris, France, and Natixis Asset Management, firstname.lastname [at] polytechnique.edu

†Université Paris 1 Panthéon-Sorbonne, MSE - CES, 106 boulevard de l'hôpital, 75013 Paris, France, firstname.lastname [at] univ-paris1.fr

This kind of noise is frequently used in the literature. It corresponds to a homogeneous perturbation of each observation of the chaotic system. Equation (2) can then be rewritten as:

$$u_t = z\left(u_{t-1} - \varepsilon_{t-1}\right) + \varepsilon_t.$$

- ▷ Dynamic noise is a set of independent identically distributed random variables (ε_t) such that:

$$u_t = z\left(u_{t-1}\right) + \varepsilon_t. \tag{3}$$

This kind of noise corresponds to a propagation of small errors. The noisy perception of the image of z feeds the dynamical system.

In the present article, we consider that the noise can be described by a random variable belonging to a wide family of random variables, the symmetric alpha-stable variables, because it includes a lot of well-known random variables and forms a general framework. A symmetric alpha-stable variable X of parameters $0 < \alpha \leq 2$, $\gamma > 0$, $\mu \in \mathbb{R}$, is characterized by its characteristic function:

$$\mathbb{E}\left[e^{itX}\right] = \exp\left(i\mu t - \gamma|t|^{\alpha}\right), \tag{4}$$

where $t \in \mathbb{R}$ and i is the imaginary unit [23][1].

Since the observed attractor, z^{ε}, is different from z, we are interested in applying a transformation on z^{ε} in order to get a good estimation of z. That crucial step of de-noising can be achieved in different ways. One of them is based on the wavelet analysis and is largely used to de-noise simple signals for which the noise has a linear influence. In the framework of chaotic systems, this approach is more pioneering [13][14]. Let us now introduce the wavelet analysis and some of its properties.

We suppose that z^{ε} is a squared-integrable function: $z^{\varepsilon} \in \mathcal{L}^2(\mathbb{R})$. It is the input signal. We want to decompose it in a discrete dyadic real wavelet basis. Such a wavelet basis is a countable subset of the function space $\mathcal{L}^2(\mathbb{R})$. Each wavelet functions is defined by:

$$\psi_{j,k} : t \in \mathbb{R} \mapsto 2^{j/2}\Psi\left(2^j t - k\right),$$

where $\Psi \in \mathcal{L}^2(\mathbb{R})$ is the real mother wavelet, $j \in \mathbb{Z}$ is the resolution level and $k \in \mathbb{Z}$ is the translation parameter [20][22]. We give two simple examples of mother wavelet:

- ▷ the Haar wavelet, defined by:

$$\Psi^{Haar} : t \in \mathbb{R} \mapsto \mathbf{1}_{[0,1/2)}(t) - \mathbf{1}_{[1/2,1)}(t);$$

- ▷ the Daubechies wavelets extend the previous case and consist in families of wavelets defined on a compact support [6]. In general, there is no short formula for such wavelet functions, except for the Haar wavelet which is the simplest case of Daubechies wavelets. However, a simple cascade algorithm allows the construction of the wavelet functions by successive interpolations.

The discrete wavelet analysis consists in decomposing the function z^{ε} in the discrete wavelet basis. For $j, k \in \mathbb{Z}$, the projection of z^{ε} on the subspace generated by the vector basis $\psi_{j,k}$ is the wavelet coefficient $\hat{z}^{\varepsilon}_{j,k}$ of resolution level j and translation parameter k:

$$\hat{z}^{\varepsilon}_{j,k} = \sum_{n=1}^{N} z^{\varepsilon}(u_n)\psi_{j,k}(u_n)(u_n - u_{n-1}), \tag{5}$$

[1] The alpha-stable distribution can be described by three parameters. If we add a fourth parameter, we can obtain the whole family of alpha-stable random variables and not only the symmetric ones. We are then able to describe a wide range of random variables. Some of them have a thin tail, others are heavy-tailed: it is therefore a much richer class than the only Gaussian random variable which is usually chosen in the literature to describe a noise. In another hand, we note that a rich class of random variables, in order to describe a noise, can be obtained with generalized hyperbolic random variables, which are described by five parameters [1][2]. Some articles make the same choice as us and they deal with a representation of the noise by alpha-stable random variables [10][16][17][18]. This family of random variables is convenient for the results that follow because, in the proofs, we use characteristic functions. We do not know the probability density function of many alpha-stable random variable but we can describe precisely at least two of them: the Gaussian variable and the Cauchy variable, which are respectively a 2-stable and a 1-stable random variable.

where we have considered the discrete grid $u_1 < u_2 < ... < u_N$ of N observations and where

$$u_0 = \min(\inf(Supp(z^\varepsilon \psi_{j,k})), u_1) \tag{6}$$

is an additional discretization point.

The wavelet inverse transform consists in reconstructing the observed signal, z^ε, according to each translation and resolution level. The reconstructed signal for a given resolution level $j \in \mathbb{Z}$ is called the detail signal:

$$D_j : t \in \mathbb{R} \mapsto \sum_{k \in \mathbb{Z}} \hat{z}^\varepsilon_{j,k} \psi_{j,k}(t).$$

The whole reconstructed signal is the sum of all the details:

$$z^\varepsilon = \sum_{j \in \mathbb{Z}} D_j.$$

As our goal is to find an approximation for the true chaotic signal z using wavelet decomposition given by (5), we propose to filter the wavelet coefficients with a *thresholding* function, following [14]. The aim of that thresholding function, Φ, consists in selecting the most significant vectors $\psi_{j,k}$ of the wavelet basis and to eliminate the too noisy terms of the wavelet decomposition. Therefore, for a wavelet coefficient $w \in \mathbb{R}$, Φ is defined by:

$$\Phi(w) = w \mathbf{1}_{\{|w| \geq \lambda\}},$$

like in [7], where we have set $\lambda > 0$ as the threshold. Then, we define $\tilde{z}$, the filtered wavelet inverse transform of the noisy attractor, by:

$$\tilde{z} : t \in \mathbb{R} \mapsto \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \Phi\left(\hat{z}^\varepsilon_{j,k}\right) \psi_{j,k}(t). \tag{7}$$

For a given $\lambda > 0$, the filtered function $\tilde{z}$, defined by equation (7), is an approximation of z. The aim of the deconvolution with the wavelet method is to select the λ which allows to get the optimal approximation $\tilde{z}$ of z, in the sense that the selected λ is the value which minimizes $d(\tilde{z}, z)$, where d is a function determining the error made by our approximation:

$$d(\tilde{z}, z) = \mathbb{E}\left[\int_{\mathbb{R}} (\tilde{z}(u) - z(u))^2 \, du\right]. \tag{8}$$

In [9], we propose such an optimal threshold. In order to calculate it rigorously, we need to know the probability density function of each wavelet coefficient $\hat{z}^\varepsilon_{j,k}$ of the noisy attractor. Indeed, such probability density functions are fundamental when we calculate the expected value d in (8).

Therefore, the present article aims to determine the probability density of each wavelet coefficients $\hat{z}^\varepsilon_{j,k}$. The importance of such a study for signal processing already arose for simpler frameworks. For example, in [24], Schoukens and Renneboog studied the probability density of each Fourier coefficients of a linear noisy signal, where the noise is represented by Gaussian random variables.

The article is divided into two parts, describing successively the theory (Section 2) and some applications (Section 3).

2 Main results

In the present section we provide the probability density function of a wavelet coefficient $\hat{z}^\varepsilon_{j,k}$ defined in equation (5) in the presence of both kinds of noise defined respectively in equations (3) and (2), namely a dynamic noise and a measurement noise. The proofs of the following theorems and propositions are reported at the end of the article.

We consider a sequence of $N \in \mathbb{N}$ noisy observations: $u_1, u_2, ..., u_N$, that we rank: $u_{1:N} \leq u_{2:N}, ... \leq u_{N:N}$. For convenience, in the following, we note directly: $u_1 \leq u_2, ... \leq u_N$.

We introduce now some assumptions cited in the following theorems and propositions:

(A0) *z is a real function with a bounded support and $j, k \in \mathbb{Z}$.*

(A1) *$u_1 \leq u_2 \leq ... \leq u_N \in \mathbb{R}$ and u_0 is such that $u_0 = \min(\inf(Supp(z^{\varepsilon}\psi_{j,k})), u_1)$.*

2.1 Dynamic noise

We begin with the case of a dynamic noise, introduced in equation (3). We introduce some assumptions for this dynamic noise:

(A2) *$\varepsilon_1, \varepsilon_2, ..., \varepsilon_N$ are N independent identically distributed symmetric alpha-stable random variables of parameters α, γ, μ, where $0 < \alpha \leq 2$, $\gamma > 0$, $\mu \in \mathbb{R}$.*

(A2g) *$\varepsilon_1, \varepsilon_2, ..., \varepsilon_N$ are N independent identically distributed centred Gaussian random variables of variance σ^2, with $\sigma > 0$.*

(A2c) *$\varepsilon_1, \varepsilon_2, ..., \varepsilon_N$ are N independent identically distributed Cauchy random variables with scale parameter γ and location parameter 0, with $\gamma > 0$.*

(A3) *The noisy and observed attractor is defined by:*

$$\forall n \in \{1, ..., N\},\ z^{\varepsilon}(u_n) = z(u_n) + \varepsilon_n.$$

We present, in Theorem 1, a general expression of the random wavelet coefficient of a chaos perturbed by alpha-stable dynamic noise. Then, we study two particular cases according to the nature of the alpha-stable noise: the Gaussian case is studied in Proposition 1 and the Cauchy noise in Proposition 2.

Theorem 1. *Let assume (A0), (A1), (A2) and (A3). Then, the wavelet coefficient $\hat{z}^{\varepsilon}_{j,k}$ of the noisy and observed attractor, z^{ε}, is a symmetric alpha-stable random variable of parameters α, γ', μ', where:*

$$\begin{cases} \gamma' &= \gamma \sum_{n=1}^{N} |\psi_{j,k}(u_n)(u_n - u_{n-1})|^{\alpha} \\ \mu' &= \hat{z}_{j,k} + \mu \sum_{n=1}^{N} \psi_{j,k}(u_n)(u_n - u_{n-1}). \end{cases}$$

If we are facing a centred Gaussian noise, then every wavelet coefficient of the noisy signal is a Gaussian variable:

Proposition 1. *Let assume (A0), (A1), (A2g) and (A3). The wavelet coefficient $\hat{z}^{\varepsilon}_{j,k}$ of the noisy observed attractor, $z^{\varepsilon}(u_n)$, is a Gaussian random variable with mean $\hat{z}_{j,k}$ and variance $(\sigma')^2$, where:*

$$\sigma' = \sigma \sqrt{\sum_{n=1}^{N} |\psi_{j,k}(u_n)(u_n - u_{n-1})|^2}.$$

Moreover, if $\psi_{j,k}$ is an orthonormal wavelet, then $\sigma' \leq \sigma$ when $N \to \infty$.

If the noise is described by a Cauchy distribution, then the wavelet coefficients of the noisy signal are also Cauchy variables. This kind of noise is interesting because its fat tail is a good way to model the real perturbations of a signal. The Cauchy distribution example is developed in Proposition 2. We call respectively location parameter and scale parameter the numbers $\mu \in \mathbb{R}$ and $\gamma > 0$ such that the Cauchy density function is:

$$\delta_{Cauchy} : x \in \mathbb{R} \mapsto \frac{1}{\pi\gamma \left[1 + \left(\frac{x-\mu}{\gamma}\right)^2\right]}.$$

For simplicity, in this article, we only consider Cauchy variables for which $\mu = 0$.

Proposition 2. *Let assume (A0), (A1), (A2c) and (A3). The wavelet coefficient $\hat{z}^{\varepsilon}_{j,k}$ of the noisy observed attractor, $z^{\varepsilon}(u_n)$, is a Cauchy random variable with location parameter $\hat{z}_{j,k}$ and scale parameter γ', where:*

$$\gamma' = \gamma \sum_{n=1}^{N} |\psi_{j,k}(u_n)(u_n - u_{n-1})|.$$

Proposition 2 is a direct consequence of Theorem 1 since any Cauchy random variable with scale parameter γ and location parameter 0 is also a symmetric 1-stable random variable with parameters $\mu = 0$ and γ.

One may be surprised by the simplicity of the results obtained for a dynamic noise. The astonishment will grow when one will read in the next section that it is more complicated to obtain de probability density function of the wavelet coefficients in a measurement noise environment. Indeed, in the dynamic noise case, the noisy attractor and the pure one are linked by an affine relation, whereas this link is much intricate in the measurement noise case. Therefore, we remark in [9] that it is easier to approximate the attractor by wavelet shrinkage when the noise is dynamic than when it is a measurement noise. On the contrary, it seems obvious that it is more complicated to predict the trajectory of the pure chaos when we are facing a dynamic noise than when we are facing a measurement one.

2.2 Measurement noise

We now assume that a measurement noise is perturbing the system, like in equation (2). We complete our list of assumptions with the following ones, specific to a measurement noise:

(A2*) *$\varepsilon_1, \varepsilon_2, ..., \varepsilon_N, \varepsilon_1^\star, \varepsilon_2^\star, ..., \varepsilon_N^\star$ are $2N$ independent identically distributed symmetric alpha-stable random variables of parameters α, γ, μ, where $0 < \alpha \leq 2$, $\gamma > 0$, $\mu \in \mathbb{R}$.*

(A2*g) *$\varepsilon_1, \varepsilon_2, ..., \varepsilon_N, \varepsilon_1^\star, \varepsilon_2^\star, ..., \varepsilon_N^\star$ are $2N$ independent identically distributed centred Gaussian random variables of variance σ^2, where $\sigma > 0$.*

(A2*c) *$\varepsilon_1, \varepsilon_2, ..., \varepsilon_N, \varepsilon_1^\star, \varepsilon_2^\star, ..., \varepsilon_N^\star$ are $2N$ independent identically distributed Cauchy random variables with scale parameter γ and location parameter 0, where $\gamma > 0$.*

(A4) *The noise is a measurement noise described by equation (3). More precisely, the noisy and observed attractor is defined by:*

$$\forall n \in \{1, ..., N\},\ z^\varepsilon(u_n) = z(u_n - \varepsilon_n^\star) + \varepsilon_n. \tag{9}$$

(A5) *z is a continuous and piecewise continuously differentiable function on $\mathbb{R}$: $\exists M \in \mathbb{N}$, $\exists \chi_0 < ... < \chi_M \in \mathbb{R}$, such that $Supp(z) \subset [\chi_0, \chi_M]$ and, $\forall m \in \{1, ..., M\}$, the restriction of z to the interval (χ_{m-1}, χ_m) is differentiable with continuous derivative and $z|'_{(\chi_{m-1},\chi_m)}$ has a limit on its right in χ_{m-1} and a limit on its left in χ_m. Moreover $\mathcal{N}$ is the set of the non-differentiable points of z and $u_1, ..., u_N \in \mathbb{R} \setminus \mathcal{N}$.*

(A6) *We only observe states of the dynamical system separated by, at least, two time steps.*

(A7) *$\max_{n \in \{1,...,N\}} |\varepsilon_n^\star|$ is small*[2].

We have to pay attention to the fact that, for an observation time t, the observation u_t depends both on ε_t and ε_{t-1}, whereas at the next observation time, $t+1$, the observation u_{t+1} depends both on ε_{t+1} and ε_t. Then, for the next observation time, $t+2$, the observation u_{t+1} depends both on ε_{t+2} and ε_{t+1}. Therefore, the independence of the noise implies the independence between u_t and u_{t+2}, but consecutive observations, such as u_t and u_{t+1}, are not independent. In order to avoid such a situation where some noises are correlated, we assume that we do not observe two consecutive states of the system[3]. But, for convenience, we simply

[2] Theoretically, the adjective *small* is clarified when we describe, in Theorem 3, the error of the approximation of Theorem 2. More practically, the extreme value theory allows to link $\max_{n \in \{1,...,N\}} |\varepsilon_n^\star|$ to the scale parameter γ of the noise. For instance, in the example about the logistic map, the error, expressed as the value of the integral of the difference, in absolute value, between the empirical probability density function and the one obtained by the application of Proposition 3, is simply 4.7% for $\sigma = 5\%$. If $\sigma = 10\%$, the error becomes about 10.9%.

[3] Formally, for an initial observation $u \in \mathbb{R}$, we only observe $\left(z^{\nu(n)}\right)_{1 \leq n \leq N-1}$, where $z^n(u)$ denotes $z^{n-1}(z(u))$ and $z^1 = z$ and where ν is a function $\mathbb{N} \to \mathbb{N}$, such that:

$$\forall n \in \mathbb{N},\ \nu(n+1) - \nu(n) \geq 2.$$

note $u_1 \leq u_2, ... \leq u_N$ these non-consecutive and ranked observations. That condition allows us to assume that $\varepsilon_1, \varepsilon_2, ..., \varepsilon_N, \varepsilon_1^\star, \varepsilon_2^\star, ..., \varepsilon_N^\star$ are independent random variables.

We now present two different approximations of the probability density function of a wavelet coefficient of a chaos with measurement noise. The first one is based on the study of the sensitivity of such a coefficient to small noise. It is particularly suitable when the noise is small. The second method is based on Edgeworth expansion and provides even better results as the number of observation points increases.

2.2.1 Approximation of the noise influence

In this paragraph, we propose first an approximation of the wavelet coefficient distribution of the noisy attractor. This approximation is detailed in Theorem 2 and Propositions 3 and 4. Then, we give an estimation of the error of such an approximation in Theorem 3.

We consider that z is continuous and piecewise continuously differentiable on $\mathbb{R}$, what is the assumption *(A5)*. In fact, such an assumption provides for z a derivative function with a finite limit on the left and the right of each point. Therefore, a first-order Taylor expansion can be achieved because the sensitivity of $z(u)$ to a small variation on u is well defined. That small variation is a consequence of the assumption *(A7)* made for the noise.

Theorem 2. *Let assume (A0), (A1), (A2***), (A4), (A5), (A6) and (A7). Then, the wavelet coefficient* $\hat{z}^\varepsilon_{j,k}$ *of the noisy and observed attractor,* z^ε*, can be approximated by a symmetric alpha-stable random variable of parameters* α*,* γ'*,* μ'*, where:*

$$\begin{cases} \gamma' &= \gamma \sum_{n=1}^{N} (1 + |z'(u_n)|^\alpha) |\psi_{j,k}(u_n)(u_n - u_{n-1})|^\alpha \\ \mu' &= \hat{z}_{j,k} + \mu \sum_{n=1}^{N} (1 - z'(u_n)) \psi_{j,k}(u_n)(u_n - u_{n-1}). \end{cases}$$

In a practical use, we can have a good approximation thanks to Theorem 2 when we replace the condition $u_1, ..., u_N \in \mathbb{R} \setminus \mathcal{N}$ by the condition: $\mathcal{N}$ is negligible for the Lebesgue measure, N is big, z is the attractor of an ergodic chaos and its invariant measure has no accumulation point in $\mathcal{N}$. Indeed, the probability of observing a point u_n in $\mathcal{N}$ would be negligible. The same remark can be done for both following propositions, namely Proposition 3 and Proposition 4.

We assume, that the noise is small *(A7)*. In practice, the random noise is unknown since we just observe the noisy data without viewing the clean signal. However, assuming we know the probability density function of the noise, we can reasonably use the results of Theorem 2 as an approximation, when such a probability density function is much localized around zero. It will be the case when the variance is small. But when the variance is not defined, other parameters may be considered, still as an approximation, like the scale parameter for a Cauchy distribution.

If we are facing a centred Gaussian noise, then the wavelet coefficient of the noisy signal is approximated by a Gaussian variable. This is the object of Proposition 3.

Proposition 3. *Let assume (A0), (A1), (A2***g), (A4), (A5), (A6) and (A7). Then, the limit of the wavelet coefficient* $\hat{z}^\varepsilon_{j,k}$ *of the noisy and observed attractor,* z^ε*, can be approximated by a Gaussian random variable with mean* $\hat{z}_{j,k}$ *and variance* $(\sigma')^2$*, where:*

$$\sigma' = \sigma \sqrt{\sum_{n=1}^{N} (1 + |z'(u_n)|^2) \, |\psi_{j,k}(u_n)(u_n - u_{n-1})|^2}.$$

Proposition 3 is a direct consequence of Theorem 2 since any centred Gaussian random variable of variance σ^2 is also a symmetric 2-stable random variable with parameters $\mu = 0$ and $\gamma = \sigma^2/2$.

If the noise is described by a Cauchy distribution, then the wavelet coefficient of the noisy signal is approximated by a Cauchy variable. This fat-tailed distribution case is the object of Proposition 4.

Proposition 4. *Let assume (A0), (A1), (A2***c), (A4), (A5), (A6) and (A7). Then, the limit of the wavelet coefficient* $\hat{z}^\varepsilon_{j,k}$ *of the noisy and observed attractor,* z^ε*, can be approximated by* *a Cauchy random variable with location parameter $\hat{z}_{j,k}$ and scale parameter γ', where:*

$$\gamma' = \gamma \sum_{n=1}^{N} (1 + |z'(u_n)|) \, |\psi_{j,k}(u_n)(u_n - u_{n-1})|.$$

Proposition 4 is a direct consequence of Theorem 2 since any Cauchy random variable with scale parameter γ and location parameter 0 is also a symmetric 1-stable random variable with parameters $\mu = 0$ and γ.

We see in Proposition 3 and Proposition 4 that the probability density functions of the resulting wavelet coefficients are more spread than in Proposition 1 and Proposition 2 around the wavelet coefficients of the signal without any perturbation by some noise. Indeed, a coefficient $|z'(u_n)|^2$ or $|z'(u_n)|$ appears in the expression of the variance or of the scale parameter. This is intuitive since the noise only linearly affects z in Proposition 1 and Proposition 2, whereas, in addition to that linear effect, a non-linear noise influence is added in Proposition 3 and Proposition 4.

It can also be interesting to have an idea of the accuracy of the approximation presented in Theorem 2. The error made achieving a Taylor expansion can be controlled. This is the aim of the end of the present paragraph. In Theorem 3, we highlight an upper bound of any quantile of the error made by the approximation suggested in Theorem 2. We restrict here a little the framework to two times continuously differentiable attractors what allow us to have an upper bound of the error thanks to Taylor expansion. Nevertheless, that error given by Taylor expansion is a random variable. In order to get a deterministic upper bound of the error, we will then use extreme value theory. Indeed, the distribution function of the random variables considered is not bounded so that, theoretically, the error can be infinite. However, an upper bound of high quantiles gives a good idea of the error and extreme value theory can furnish such quantiles.

The error estimation, and more particularly the use of extreme value theory, incites us to add two assumptions:

(A8) *g_N is the cumulative distribution function of $\sum_{n=1}^{N} |\varepsilon_n^\star|^2$.*

(A9) *The cumulative distribution function of $\left(\max_{n \in \{1,\dots,N\}} \left(|\varepsilon_n^\star|^2\right) \right)$ when N is big exists and is denoted h_N.*

Theorem 3. *Let assume (A0), (A1), (A2*), (A4), (A5), (A6), (A7), (A8) and (A9). Moreover, z is assumed to be two times continuously differentiable and corresponding to the attractor of an ergodic chaos. We assume that $\mathcal{N}$ is Lebesgue measure-zero. Let $0 \leq p \leq 1$. Let Δ_U be the cumulative distribution function of the independent identically distributed observations (u_n) and let δ_U be the corresponding probability density function. Let $\hat{Z}^{\varepsilon}_{j,k}$ be the approximation[4] of the wavelet coefficient $\hat{z}_{j,k}$ defined by:*

$$\hat{Z}^{\varepsilon}_{j,k} = \hat{z}_{j,k} + \sum_{n=1}^{N} \left[-z'(u_n)\varepsilon_n^\star + \varepsilon_n \right] \psi_{j,k}(u_n)(u_n - u_{n-1}).$$

We define a random distance $d_{j,k}$ between $\hat{Z}^{\varepsilon}_{j,k}$ and $\hat{z}^{\varepsilon}_{j,k}$ by:

$$d_{j,k} = \left| \hat{z}^{\varepsilon}_{j,k} - \hat{Z}^{\varepsilon}_{j,k} \right|.$$

We define $D_{j,k}(p)$, the quantiles of $d_{j,k}$, by $\mathbb{P}\left[d_{j,k} \leq D_{j,k}(p)\right] = p$. Then, there exists $\Omega_N, \Omega_N^\infty$: $[0,1] \to \mathbb{R}$ such that these quantiles $D_{j,k}(p)$ have an upper bound:

$$D_{j,k}(p) \leq \min\left(\Omega_N(p), \Omega_N^\infty(p)\right),$$

[4] It is the approximation appearing in equation (13) in the proof of Theorem 2

where

$$\begin{cases} \Omega_N(p) & = & \left[\frac{1}{2}\max\limits_{x\in Supp(z)}\left(|z''(x)|\right)\max\limits_{x\in Supp(\psi_{j,k})}\left(|\psi_{j,k}(x)|\right)\right] f_N^{-1}(\sqrt{p})g_N^{-1}(\sqrt{p}) \\ \Omega_N^{\infty}(p) & \overset{N\to\infty}{\sim} & \left[\frac{1}{2}\max\limits_{x\in Supp(z)}\left(|z''(x)|\right)\int_{\mathbb{R}}|\psi_{j,k}(x)|\,dx\right] h_N^{-1}(p) \end{cases}$$

and where

$$f_N : v>0 \mapsto \max\left\{0, 1-N\int_{\inf(Supp(\delta_U))+v}^{\sup(Supp(\delta_U))}\delta_U(u)\left[\Delta_U(u-v)+1-\Delta_U(u)\right]^{N-1}du\right\}.$$

Theorem 3 gives an upper bound for every quantile of the error made in the approximation of the wavelet coefficient of a chaos with a measurement noise as proposed in Theorem 2. It is therefore a useful tool in order to have a wise and prudent estimation of the accuracy of our approximation since it exaggerates the error. Indeed, the probability that $d_{j,k}$ is bellow $\Omega_N(p)$ or $\Omega_N^{\infty}(p)$ is greater than p. Moreover, it introduces two kinds of upper bounds, one when the number of observations N is finite and another one when that number tends towards infinity. From a practical point of view, Theorem 3 may be more explicit about g_N and h_N if we are facing a Gaussian noise. This is the aim of Proposition 5. We will further consider the case of a Cauchy noise in Proposition 6.

Proposition 5. *Let $x \geq 0$. Let assume (A2*g), (A8) and (A9). Then:*

$$\begin{cases} g_N(x) & = & \dfrac{\gamma\left(\frac{N}{2},\frac{\sigma^2 x}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} \\ h_N(x) & \overset{N\to\infty}{\sim} & \exp\left(-\exp\left(\alpha(N)-\frac{x}{2\sigma^2}\right)\right), \end{cases}$$

where $\alpha(N)$ is such that:

$$\alpha(N)^{\frac{1}{2}}e^{\alpha(N)} = \frac{N}{\sqrt{\pi}}$$

and Γ and γ are respectively the Gamma function and the lower incomplete Gamma function[5]*.*

If the noise is a Cauchy random variable, then we can achieve the same work as in Proposition 5 in order to make some functions of Theorem 3 more explicit. This is the aim of Proposition 6. This result only highlights the case $N \to +\infty$.

Proposition 6. *Let $x \geq 0$. Let assume (A2*c) and (A9). Then:*

$$h_N(x) \overset{N\to\infty}{\sim} \exp\left(-\frac{2\gamma N}{\pi\sqrt{x}}\right).$$

2.2.2 Expansion of the probability density

In the approximation of the noise influence in Theorem 2, we limit our analysis to a local impact of the noise. Indeed, we perform a linear approximation of this influence at a given point of the support of the attractor function. Implicitly, we restrict ourself to a small noise. We could ameliorate our analysis by achieving a second order Taylor expansion, but such a result would still consist on a local analysis of the noise influence. In order to have a global view and to take into account a larger noise, we can realize an expansion of the probability density function. With that approach, the convergence of our estimator is much less dependent on the size of the noise but more on the size of the observation sample.

Then, we use the well-known Edgeworth expansion as in [3][5][19] with the notations adapted to our framework. Let assume *(A0)*, *(A1)*, *(A2*)*, *(A4)*, *(A5)*, *(A6)*. We assume moreover that all the moments of the $2N$ random variables representing the noise are well

[5] For $x, k \geq 0$:

$$\begin{cases} \Gamma(k) & = & \int_0^{+\infty} t^{k-1}e^{-t}dt \\ \gamma(k,x) & = & \int_0^{x} t^{k-1}e^{-t}dt. \end{cases}$$

defined. Let ϕ_N be the probability density function of a random Gaussian variable with mean μ_N and variance σ_N^2, where:

$$\begin{cases} \sigma_N^2 & = & \gamma \sum_{n=1}^N (1+|z'(u_n)|^\alpha)|\psi_{j,k}(u_n)(u_n-u_{n-1})|^\alpha \\ \mu_N & = & \hat{z}_{j,k} + \mu \sum_{n=1}^N (1-z'(u_n))\psi_{j,k}(u_n)(u_n-u_{n-1}). \end{cases}$$

We note Φ_N the corresponding cumulative distribution function. If $\sigma_N^2 < 2$, then the probability density function, $\delta_{\hat{z}^\varepsilon_{j,k}}$, of the wavelet coefficient $\hat{z}^\varepsilon_{j,k}$ of the noisy and observed attractor is such that, for $x \in \mathbb{R}$:

$$\delta_{\hat{z}^\varepsilon_{j,k}}(x) = \phi_N(x)\left[1 + \sum_{m=1}^{+\infty} \mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu_N}{\sqrt{2}\sigma_N}\right)\right]\mathcal{H}_m\left(\frac{x-\mu_N}{\sqrt{2}\sigma_N}\right)\frac{\sigma_N^m}{(\sqrt{2})^m m!}\right], \qquad (10)$$

where $(\mathcal{H}_m)_{m\in\mathbb{N}}$ are Hermite polynomials[6]. The corresponding cumulative distribution function, $\Delta_{\hat{z}^\varepsilon_{j,k}}$, of the wavelet coefficient $\hat{z}^\varepsilon_{j,k}$ is such that, for $x \in \mathbb{R}$:

$$\Delta_{\hat{z}^\varepsilon_{j,k}}(x) = \Phi_N(x) + \sum_{m=1}^{+\infty} \mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu_N}{\sqrt{2}\sigma_N}\right)\right]\left[\frac{\mathcal{H}_{m-1}(0)}{\sqrt{2}\sigma_N} - \phi_N(x)\mathcal{H}_{m-1}\left(\frac{x-\mu_N}{\sqrt{2}\sigma_N}\right)\right]\frac{\sigma_N^{m+1}}{(\sqrt{2})^{m-1}m!}. \qquad (11)$$

In a practical use of equation (10), we restrict to truncations of the expansion. We have to pay attention to the fact that such truncated expressions do not strictly define a probability density function and a cumulative distribution function. Therefore, the result will be even more reliable for high-order truncations. It is then recommended to set a stopping criterion and to truncate the sum in equation (10) when new terms seem negligible.

In equation (10), because of the Hermite polynomials, we implicitly need the expression of the moments of the wavelet coefficient of the noisy attractor. These moments cannot be deducted directly from the moments of the attractor value in u_n, even though we get the wavelet coefficient from a simple linear combination of $z^\varepsilon(u_1), ..., z^\varepsilon(u_N)$. Instead of that, we propose to calculate jointly the moments and the cumulants. Indeed, cumulants seem to be an appropriate tool[7] and they incite us to propose the following algorithm in order to get the value of $\mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu_N}{\sqrt{2}\sigma_N}\right)\right]$ in equation (10) determining the probability density expansion of the wavelet coefficients. More precisely, we are going to define two algorithms:

- ▷ in Algorithm 1, we define a procedure which allows to calculate $\mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu_N}{\sqrt{2}\sigma_N}\right)\right]$ from the cumulants of $z^\varepsilon(u_n)$;
- ▷ considering that the cumulants of $z^\varepsilon(u_n)$ are not necessarily easy to get, we define in Algorithm 2 a way to get them from the cumulants of the noise in the case where z can be written as a convergent power series.

The reading of these algorithms is somewhat laborious, but their implementation is very easy.

[6] $\forall (m,x) \in \mathbb{N}\times\mathbb{R},\ \mathcal{H}_m(x) = (-1)^m e^{x^2/2}\frac{d^m}{dx^m}e^{-x^2/2}$.

[7] Indeed, it is easy to get moments from cumulants and cumulants from moments thanks to the fact that $C_1(X) = M_1(X)$ and to the recursive formula, for $r \in \{1,2,...\}$:

$$C_r(X) = M_r(X) - \sum_{s=1}^{r-1}\binom{r}{s}C_s(X)M_{r-s}(X),$$

where C denotes any finite cumulant, M any finite moment and X a random variable for which cumulants and moments are well defined. Moreover, two properties concerning cumulants are very interesting:

- ▷ The cumulant of any sum of independent random variables is the sum of the cumulants of each random variable.
- ▷ The cumulants are homogeneous: the r-th cumulant of aX, for any $a \in \mathbb{R}$, is $a^r C_r(X)$.

Algorithm 1. *For $r \in \{1, 2, ...\}$, $C_r(X)$ and $M_r(X)$ are respectively the r-th cumulant and the r-th moment of the random variable X for which cumulants and moments are well defined. We suppose that those cumulants are known for the random variables $z^\varepsilon(u_n)$, where $n \in \{1, ..., N\}$. Let $M \in \mathbb{N}$ and $\mu, \sigma \in \mathbb{R}$. Then, the following algorithm allows to calculate $\mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)\right]$ for every $m \in \{1, ..., M\}$, where $\mathcal{H}_m$ is the m-th Hermite polynomial:*

1. *We begin with $r = 1$. At each iteration, we replace r by $r + 1$.*
2. *For each j and k indexing the considered wavelet coefficients, we calculate $C_r\left(\hat{z}^\varepsilon_{j,k}\right)$ as:*
$$C_r\left(\hat{z}^\varepsilon_{j,k}\right) = \sum_{n=1}^{N} \left[\psi_{j,k}(u_n)(u_n - u_{n-1})\right]^r C_r\left(z^\varepsilon(u_n)\right).$$
3. *For each j and k indexing the considered wavelet coefficients, we calculate $C_r\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)$ as:*
$$C_r\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) = \frac{C_r\left(\hat{z}^\varepsilon_{j,k}\right)}{\left[\sqrt{2}\sigma\right]^r} - \frac{C_r(\mu)}{\left[\sqrt{2}\sigma\right]^r}.$$
4. *For each j and k indexing the considered wavelet coefficients, we calculate $M_r\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)$ as:*
$$\begin{cases} M_r\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) = C_r\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) & \text{if } r = 1 \\ M_r\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) = C_r\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) + \sum_{s=1}^{r-1}\binom{r}{s} C_s\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) M_{r-s}\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) & \text{else.} \end{cases}$$
5. *We go back to the step 1 of the algorithm until $r = M$ is reached.*
6. *For each $m \in \{0, ..., M\}$, we deduct $\mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)\right]$ from the previously calculated moments, say:*
$$\begin{cases} \mathbb{E}\left[\mathcal{H}_0\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)\right] = 1 \\ \mathbb{E}\left[\mathcal{H}_1\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)\right] = M_1\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) \\ \mathbb{E}\left[\mathcal{H}_2\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)\right] = M_2\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) - 1 \\ \mathbb{E}\left[\mathcal{H}_3\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)\right] = M_3\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) - 3M_1\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right) \end{cases}$$
and so on until the M-th polynomial.

The degree of the m-th order Hermite polynomial is precisely m. Therefore, we only need the cumulants until the order m in order to calculate $\mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}^\varepsilon_{j,k}-\mu}{\sqrt{2}\sigma}\right)\right]$. That consideration justify the fifth step of Algorithm 1. We also note that the algorithmic complexity of Algorithm 1, when the size of the sample grows towards infinity, is dominated[8] by $M(N + \xi)$, where ξ is the number of wavelet coefficients considered, depending of course on the number of observed data, N.

Algorithm 2 calculates the cumulants of $z^\varepsilon(u_n)$ from the cumulants of the noise in the case where z can be written as a convergent power series.

Algorithm 2. *Let z be a function on $\mathbb{R}$ which can be written as a convergent power series around a center c on its support: for each $x \in Supp(z)$, $\exists(a_k)_{k\in\mathbb{N}}$ such that $z(x) = \sum_{k=0}^{+\infty} a_k(x-c)^k$. Let $K \in \mathbb{N}$ be the truncation order retained in the computation of z: $z(x) \approx \sum_{k=0}^{K} a_k(x-c)^k$. Let $\varepsilon^\star_1, \varepsilon^\star_2, ..., \varepsilon^\star_N, \varepsilon_1, \varepsilon_2, ..., \varepsilon_N$ be $2N$ independent identically distributed random variables with all their cumulants well defined. For $r \in \{1, 2, ...\}$, $C_r(X)$ and $M_r(X)$ are respectively the r-th cumulant and the r-th moment of the random variable X for which cumulants and moments are well defined. Let $R \in \mathbb{N}$. Then, the following algorithm allows to calculate the R first cumulants of $z^\varepsilon(u_n)$, where $n \in \{1, ..., N\}$ and $z^\varepsilon(u_n) = z(u_n - \varepsilon^\star_n) + \varepsilon_n$:*

[8] If $M = 10$, $N = 1000$ and $\xi < N$, the execution time is about 0.5 second for a classical interpreted language running on an Intel Core 2 Duo 3GHz.

1. *For each $n \in \{1, ..., N\}$ and each $r \in \{1, ..., KR\}$, we calculate $M_r(u_n - c - \varepsilon_n^\star)$ as:*

$$M_r(u_n - c - \varepsilon_n^\star) = C_r(u_n - c) - C_r(\varepsilon_n^\star) + \sum_{s=1}^{r-1} \binom{r}{s} [C_s(u_n - c) - C_s(\varepsilon_n^\star)] M_{r-s}(u_n - c - \varepsilon_n^\star).$$

2. *For each $n \in \{1, ..., N\}$, each $r \in \{1, ..., R\}$ and each $k \in \{1, ..., K\}$ we calculate $C_r\left((u_n - c - \varepsilon_n^\star)^k\right)$ as:*

$$C_r\left((u_n - c - \varepsilon_n^\star)^k\right) = M_{kr}(u_n - c - \varepsilon_n^\star) - \sum_{s=1}^{r-1} \binom{r}{s} C_s((u_n - c - \varepsilon_n^\star)^k) M_{k(r-s)}(u_n - c - \varepsilon_n^\star).$$

3. *For each $n \in \{1, ..., N\}$ and $r \in \{1, ..., R\}$, we calculate the cumulant of $z^\varepsilon(u_n)$ as:*

$$C_r(z^\varepsilon(u_n)) \approx C_r(\varepsilon_n) + \sum_{k=0}^{K} a_k^r C_r\left((u_n - c - \varepsilon_n^\star)^k\right),$$

where the approximation stands for the truncation at the order K in the power series.

We can then use successively Algorithm 2 and Algorithm 1 in order to estimate $\mathbb{E}\left[\mathcal{H}_m\left(\frac{\hat{z}_{j,k}^\varepsilon - \mu_N}{\sqrt{2}\sigma_N}\right)\right]$ in the case of a convergent power series. The algorithmic complexity of Algorithm 2, when the size of the sample grows towards infinity, is dominated[9] by $(KR)^2 N$. However, that algorithmic cost may be strongly reduced when the noise is a Gaussian variable, while remaining linear in N. Indeed, only two cumulants are then different from zero.

We note that such an approach does not work for every kind of symmetric stable noise because cumulants have to be defined. This is the case, for example, for Gaussian noise but not for Cauchy noise.

We also note that few maps are convergent power series. However, if z is simply a continuous function, then, according to Weierstrass Theorem, it can be uniformly approximated by a polynomial function and therefore by a power series. In practice, one can sometimes get an accurate polynomial function by Lagrange interpolation: the interpolated polynomial is equal to z at pre-specified nodes. If those nodes are equally spaced, one can face Runge's phenomenon: the interpolation is not always better when the number of interpolation nodes increases. To avoid such a *polynomial wiggle*, one may use Chebyshev nodes. Then, such interpolation polynomials often converge towards the interpolated function when then number of nodes increases [11][21][25]. Therefore, Algorithm 1 and Algorithm 2 allow to apply equation (10) and equation (11) in many practical cases, where z is simply continuous.

2.3 Extension to multidimensional chaos

Until this section, we have considered unidimensional chaos. We used therefore unidimensional wavelet basis. In practice we can often face multidimensional chaos. This implies the use of multidimensional wavelet functions. In this new framework, we denote an ergodic chaotic system by the vector $X_t = (x_{1,t}, ..., x_{p,t})$, where $p > 1$ is the dimension of the chaos. The attractor of X is a function $Z : \mathbb{R}^p \to \mathbb{R}^p$ with a bounded support, such that:

$$X_t = Z(X_{t-1}) = (z_1(x_{1,t-1}, ..., x_{p,t-1}), ..., z_p(x_{1,t-1}, ..., x_{p,t-1})).$$

Subsequently, we define the noisy attractor, Z^ε, of the noisy chaos, U_t:

$$U_t = Z^\varepsilon(U_{t-1}) = \left(z_1^\varepsilon(u_{1,t-1}, ..., u_{p,t-1}), ..., z_p^\varepsilon(u_{1,t-1}, ..., u_{p,t-1})\right).$$

After having described the multidimensional chaos, we have to define the multidimensional wavelet function. We consider that ϕ is a scaling function and that Ψ is the corresponding mother wavelet generating, by translation and dilatation, an orthonormal basis of the function

[9] If $K = 10$, $R = 10$ and $N = 1000$, the execution time is about 45 seconds for a classical interpreted language running on an Intel Core 2 Duo 3GHz.

space $\mathcal{L}^2(\mathbb{R})$. Let $\theta^0 = \phi$ and $\theta^1 = \Psi$. Then, for each integer $\eta \in \{1, ..., 2^p - 1\}$ written as a binary expression $\eta = \eta_1 ... \eta_p$, we define the family of $2^p - 1$ multidimensional mother wavelets by:

$$\Psi^\eta : T = (t_1, ..., t_p) \in \mathbb{R}^p \mapsto \theta^{\eta_1}(t_1) \times ... \times \theta^{\eta_p}(t_p).$$

By dilating and translating these mother wavelets, with all the parameters $j \in \mathbb{Z}$, $k = (k_1, ..., k_p) \in \mathbb{Z}^p$ and $\eta \in \{1, ..., 2^p - 1\}$, we get an orthonormal basis of $\mathcal{L}^2(\mathbb{R}^p)$ [20]. The wavelet functions thus obtained are:

$$\psi_{j,k}^\eta : T = (t_1, ..., t_p) \in \mathbb{R}^p \mapsto 2^{-pj/2} \Psi^\eta \left(2^j t_1 - k_1, ..., 2^j t_p - k_p \right).$$

Finally, the discrete wavelet coefficient, of resolution level $j \in \mathbb{Z}$ and translation parameter $k \in \mathbb{Z}^p$, of the multidimensional noisy attractor, Z^ε, is a vector of $\mathbb{R}^p$:

$$\hat{Z}_{\eta,j,k}^\varepsilon = \left(\sum_{n=1}^N z_1^\varepsilon(U_n) \psi_{j,k}^\eta(U_n) \int_{V(U_n)} dx, ..., \sum_{n=1}^N z_p^\varepsilon(U_n) \psi_{j,k}^\eta(U_n) \int_{V(U_n)} dx \right),$$

where, for $n \in \{1, ..., N\}$, $V(U_n) = \{X \in Supp(Z) \,|\forall m \neq n, \|X - U_n\|_{\mathbb{R}^p} \leq \|X - U_m\|_{\mathbb{R}^p}\}$ is the Voronoi region of one of the $N \in \mathbb{N}$ observations, $\|.\|_{\mathbb{R}^p}$ being the Euclidean norm in $\mathbb{R}^p$. The Voronoi region is a straight way to discretize the integral. Since we had another simple expression for unidimensional chaos, we have not used Voronoi regions in dimension 1. However, the chaos being ergodic, when N tends towards infinity both discretization techniques lead to the same result.

We can directly adapt the results developed in the unidimensional environment to multidimensional chaos. Particularly, Theorem 1 and Theorem 2 are slightly transformed to become Theorem 4 and Theorem 5.

Theorem 4. *Let $0 < \alpha \leq 2$, $\gamma > 0$, $\mu \in \mathbb{R}$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^p$, $\eta \in \{1, ..., 2^p - 1\}$, $(U_1, ..., U_N) \in \mathbb{R}^N$. Let $\varepsilon_{1,1}, \varepsilon_{1,2}, ..., \varepsilon_{1,N}, ..., \varepsilon_{p,1}, \varepsilon_{p,2}, ..., \varepsilon_{p,N}$ be pN independent identically distributed symmetric alpha-stable random variables of parameters α, γ, μ. The wavelet coefficient $\hat{Z}_{\eta,j,k}^\varepsilon$ of the noisy and observed attractor,*

$$\forall n \in \{1, ..., N\}, \;\; Z^\varepsilon(U_n) = Z(U_n) + (\varepsilon_{1,n}, ..., \varepsilon_{p,n}),$$

is a vector whose coordinate indexed by $m \in \{1, ..., p\}$ is a symmetric alpha-stable random variable of parameters α, γ_m, μ_m, where:

$$\begin{cases} \gamma_m &= \gamma \sum_{n=1}^N |\psi_{j,k}^\eta(U_n) \int_{V(U_n)} dx|^\alpha \\ \mu_m &= \left(\hat{Z}_{\eta,j,k}\right)_m + \mu \sum_{n=1}^N \psi_{j,k}^\eta(U_n) \int_{V(U_n)} dx. \end{cases}$$

Theorem 5. *Let Z be a continuous and piecewise continuously differentiable function on $\mathbb{R}^p$. Let $\mathcal{N}$ be the set of the non-differentiable points of Z. Let $0 < \alpha \leq 2$, $\gamma > 0$, $\mu \in \mathbb{R}$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^p$, $\eta \in \{1, ..., 2^p - 1\}$, $(U_1, ..., U_N) \in \mathbb{R}^N \setminus \mathcal{N}$. Let $\varepsilon_{1,1}, \varepsilon_{1,2}, ..., \varepsilon_{1,N}, ..., \varepsilon_{p,1}, \varepsilon_{p,2}, ..., \varepsilon_{p,N}$ and $\varepsilon_{1,1}^\star, \varepsilon_{1,2}^\star, ..., \varepsilon_{1,N}^\star, ..., \varepsilon_{p,1}^\star, \varepsilon_{p,2}^\star, ..., \varepsilon_{p,N}^\star$ be $2pN$ independent identically distributed symmetric alpha-stable random variables of parameters α, γ, μ. The wavelet coefficient $\hat{Z}_{\eta,j,k}^\varepsilon$ of the noisy and observed attractor,*

$$\forall n \in \{1, ..., N\}, \;\; Z^\varepsilon(U_n) = Z(U_n - (\varepsilon_{1,n}^\star, ..., \varepsilon_{p,n}^\star)) + (\varepsilon_{1,n}, ..., \varepsilon_{p,n}),$$

when $\max\limits_{n \in \{1,...,N\}, m \in \{1,...,p\}} |\varepsilon_{m,n}^\star|$ is small, can be approximated by a vector whose coordinate indexed by $m \in \{1, ..., p\}$ is a symmetric alpha-stable random variable of parameters α, γ_m, μ_m, where:

$$\begin{cases} \gamma_m &= \gamma \sum_{n=1}^N \left(1 + \sum_{l=1}^p \left|(D_l Z(U_n))_m\right|^\alpha\right) |\psi_{j,k}^\eta(U_n) \int_{V(U_n)} dx|^\alpha \\ \mu_m &= \left(\hat{Z}_{\eta,j,k}\right)_m + \mu \sum_{n=1}^N \left(1 - \sum_{l=1}^p (D_l Z(U_n))_m\right) \psi_{j,k}^\eta(U_n) \int_{V(U_n)} dx, \end{cases}$$

where D_m is the differential on the m-th coordinate.

3 Examples for diverse chaos and noises

The aim of this section is to present the different results for specific cases. We will consider two simple chaos described by the logistic map and the tent map. For each case, we present different approximations of the probability density function of a wavelet coefficient of the noisy chaos and we estimate their accuracy.

3.1 The logistic map

The logistic map of parameter α is defined by the recurrence relation:

$$u_{n+1} = z(u_n) = \alpha u_n(1 - u_n).$$

If $\alpha = 4$, then $(u_n)_{n\in\mathbb{N}}$ is a chaotic system. It is a popular example of chaos due to its simple expression [4][12].

We begin with a Gaussian noise and we represent the results in the Figure 1. We have chosen a Daubechies mother wavelet with 5 vanishing moments and the wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$, whose wavelet function $\psi_{3,4}$ has its support roughly contained in the support of z. For a centred Gaussian noise with standard deviation $\sigma = 5\%$ and a sample size $N = 129$, we obtain a negligible wavelet coefficient for the pure signal: $\hat{z}_{3,4} = -3.75 \times 10^{-13}$. The impact of the noise on $\hat{z}^{\varepsilon}_{3,4}$ is therefore preponderant as we can observe in the Figure 1.

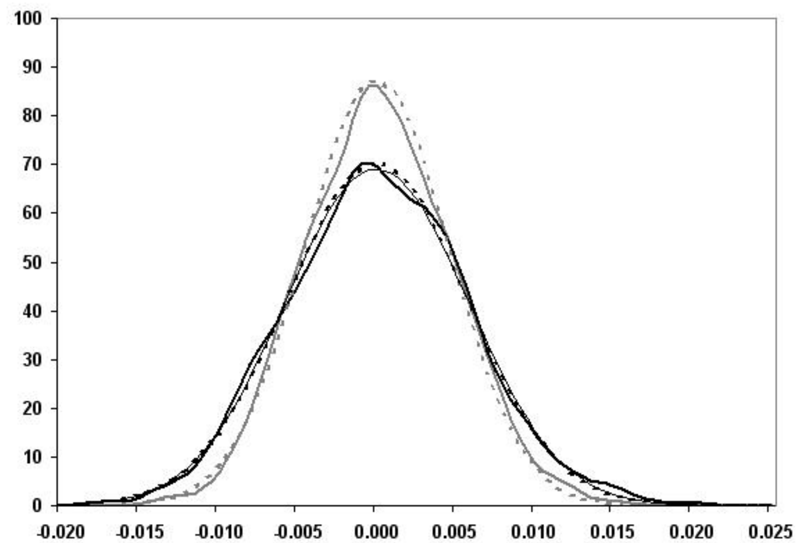

Figure 1: In grey, probability density function of the wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$ of the logistic map of parameter 4 for a centred Gaussian dynamic noise with standard deviation 0.05. If the dynamic noise is replaced by a centred Gaussian measurement noise with standard deviation 0.05, we get the black lines. In both cases, the solid thick line is the empirical density obtained by 1,000 simulations and represented with a Gaussian kernel, the dotted line is the exact density obtained by Proposition 1 for the dynamic noise and the approximation obtained by Proposition 3 for the measurement noise. The solid thin black line is the approximation of the probability density function obtained by expansion as in equation (10) truncated at the 10th order. The mother wavelet is a Daubechies wavelet with 5 vanishing moments. Moreover, the number of observations, N, is 129.

We can also see that both the proposed estimators are very good for the measurement noise as they are close to the empirical distribution. Moreover, in that case, the approximation obtained by Proposition 3, corresponding to the approximation of the noise influence, is also close to the approximation of the probability density function obtained by expansion as in equation (10). However, the second approximation is supposed to be more accurate than the first one since it is an expansion from that first approximation. In fact, there is a slight difference between both densities, represented in Figure 2: in the first approximation, the tail are slightly underestimated.

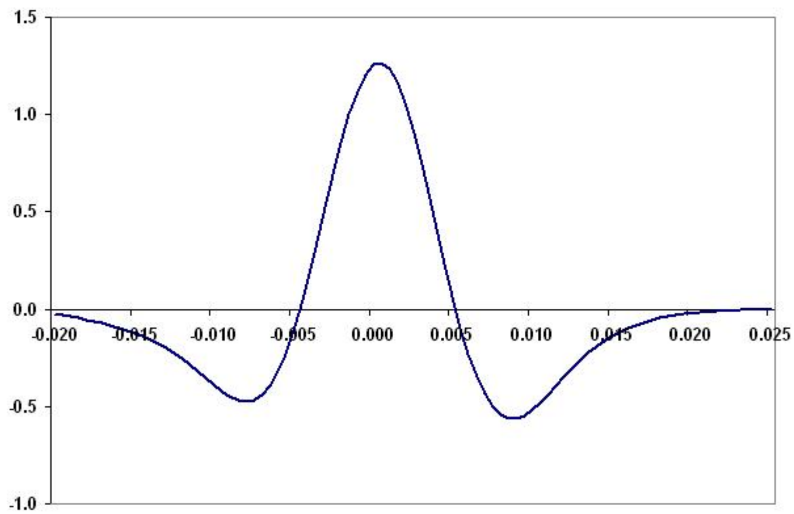

Figure 2: Difference between the approximations of the probability density functions for the measurement noise framework represented in Figure 1: the one obtained by Proposition 3 minus the one obtained by equation (10).

We can also compare, in the Figure 1, the probability density functions for the measurement noise and the dynamic noise. The first one has heavier tails than the second one, what is consistent with the fact that the noise influence is greater in the measurement noise framework.

Gaussian noise has the disadvantage to present thin tails in its probability density function, what is poorly realistic. However, the results developed in the present article allow us to deal with many kinds of noise. In particular, we present in the Figure 3 the probability density function of the wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$ when we face a Cauchy noise with scale parameter $\gamma = 1\%$.

In addition to the apparent consistency between theoretic and empirical results in the Figure 3, it seems difficult to build an empirical probability density function in a small time for Cauchy variables. Therefore, Proposition 2 and Proposition 4 provide efficient tools in order to get the exact or an approximated probability density function of the wavelet coefficients of a noisy chaos.

3.2 The tent map

The tent map of parameter α is defined by the recurrence relation:

$$u_{n+1} = z(u_n) = \begin{cases} \alpha u_n & \text{if } 0 < u_n \leq \frac{1}{2} \\ \alpha(1-u_n) & \text{if } \frac{1}{2} < u_n \leq 1 \\ 0 & \text{else.} \end{cases}$$

We consider that $\alpha = 2$. Then $(u_n)_{n\in\mathbb{N}}$ is a chaotic system [12].

We consider a Gaussian noise and we represent the results in the Figure 4. We have chosen the same Daubechies mother wavelet with 5 vanishing moments and the same wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$ than for the logistic map. For a centred Gaussian noise with standard deviation $\sigma = 5\%$ and a sample size $N = 129$, we obtain a non-negligible wavelet coefficient for the pure signal: $\hat{z}_{3,4} = -1.24 \times 10^{-2}$. The impact of the noise on $\hat{z}^{\varepsilon}_{3,4}$ is nevertheless considerable as we can observe in the Figure 1.

For the dynamic noise environment, we have a good fit between the empirical distribution of the wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$ and its exact distribution obtained by Proposition 1. The

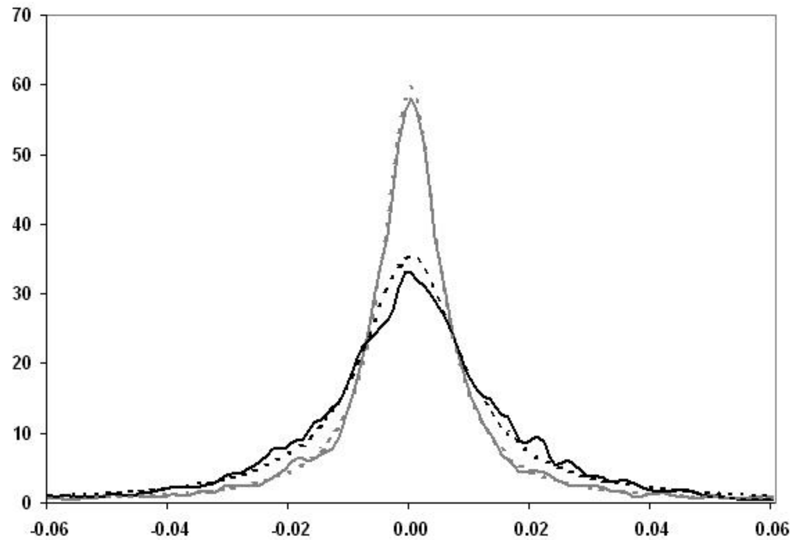

Figure 3: In grey, probability density function of the wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$ of the logistic map of parameter 4 for a Cauchy dynamic noise with scale parameter 0.01 and location parameter 0. If the dynamic noise is replaced by a Cauchy measurement noise with scale parameter 0.01 and location parameter 0, we get the black lines. In both cases, the solid thick line is the empirical density obtained by 5,000 simulations and represented with a Gaussian kernel, the dotted line is the exact density obtained by Proposition 2 for the dynamic noise and the approximation obtained by Proposition 4 for the measurement noise. The mother wavelet is a Daubechies wavelet with 5 vanishing moments. Moreover, the number of observations, N, is 129.

measurement noise case is more complex. A translation appears between the approximation obtained by Proposition 3 and the empirical distribution[10]. Thus, we must ameliorate that approximation by proposing an other method. Can it be achieved by a density expansion?

For the expansion as in equation (10), we face the problem that the tent map is not a polynomial. Unfortunately, the approximation with Chebyshev interpolation [21] seems to fail as it gives unsatisfactory results[11] strongly depending on the choice of parameters of the interpolation. In particular, we observe in Figure 4 a translation of the wavelet coefficient distribution in the measurement noise environment compared with the dynamic noise. This effect (which may be different for other wavelet coefficients) is caused by the non-differentiable point. Since the proposed polynomial interpolation does not approximate accurately z around that point, the translation effect in the approximation of the wavelet coefficient does not appear. We can therefore not use equation (10) in order to get the probability density function of the wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$.

However, we are able, for the tent map, to calculate directly the mean of any wavelet coefficient of the map with a measurement noise. Indeed, if the noise is a centred Gaussian variable of variance σ^2, then, for $j,k\in\mathbb{Z}$, ϕ_σ the Gaussian density function with variance σ^2 and Φ_σ the Gaussian cumulative distribution function with the same variance σ^2:

$$
\begin{aligned}
\mathbb{E}\left[\hat{z}^{\varepsilon}_{j,k}\right] &= \mathbb{E}\left[\sum_{n=1}^{N}\left[z(u_n-\varepsilon^{\star}_n)+\varepsilon_n\right]\psi_{j,k}(u_n)(u_n-u_{n-1})\right]\\
&= \sum_{n=1}^{N}\mathbb{E}\left[z(u_n-\varepsilon^{\star}_n)\right]\psi_{j,k}(u_n)(u_n-u_{n-1})\\
&= \sum_{n=1}^{N}\left(\int_{\mathbb{R}} z(u_n-x)\phi_\sigma(x)dx\right)\psi_{j,k}(u_n)(u_n-u_{n-1})\\
&= \sum_{n=1}^{N}\left(\int_{u_n-1}^{u_n-\frac{1}{2}}[2-2(u_n-x)]\phi_\sigma(x)dx+\int_{u_n-\frac{1}{2}}^{u_n}2(u_n-x)\phi_\sigma(x)dx\right)\psi_{j,k}(u_n)(u_n-u_{n-1})\\
&= \sum_{n=1}^{N}\Big([2-2u_n]\left[\Phi_\sigma(u_n-\tfrac{1}{2})-\Phi_\sigma(u_n-1)\right]+2u_n\left[\Phi_\sigma(u_n)-\Phi_\sigma(u_n-\tfrac{1}{2})\right]\\
&\quad +2\sigma^2\left[\phi_\sigma(u_n)-2\phi_\sigma(u_n-\tfrac{1}{2})+\phi_\sigma(u_n-1)\right]\Big)\psi_{j,k}(u_n)(u_n-u_{n-1}).
\end{aligned}
$$

[10] We remark that the only observation of that empirical distribution, which is almost centred on 0, can lead to the wrong intuition that the wavelet coefficient of the pure signal is negligible.

[11] We get a polynomial written like a truncated power series. See appendix I

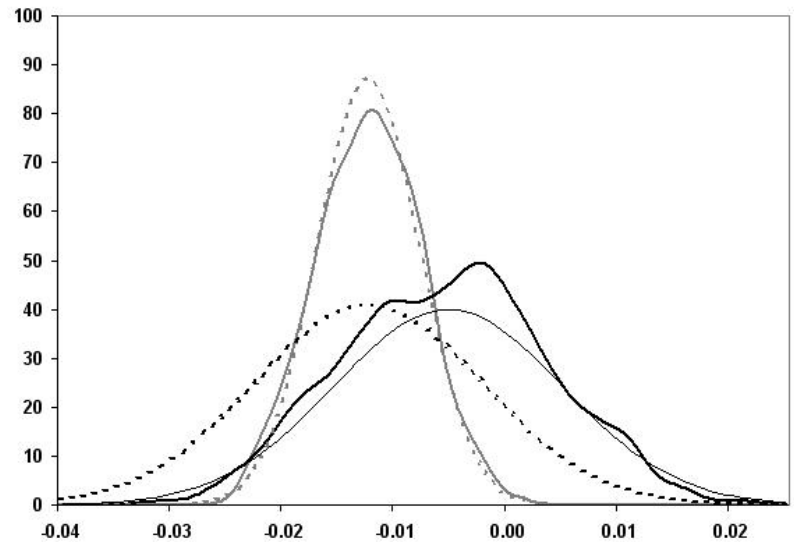

Figure 4: In grey, probability density function of the wavelet coefficient $\hat{z}^{\varepsilon}_{3,4}$ of the tent map of parameter 2 for a centred Gaussian dynamic noise with standard deviation 0.05. If the dynamic noise is replaced by a centred Gaussian measurement noise with standard deviation 0.05, we get the black lines. In both cases, the solid thick line is the empirical density obtained by 1,000 simulations and represented with a Gaussian kernel, the dotted line is the exact density obtained by Proposition 1 for the dynamic noise and the approximation obtained by Proposition 3 for the measurement noise. That approximation is ameliorated by the approximation represented with the solid thin black which is a Gaussian density with the same standard deviation as the one calculated thanks to Proposition 3 but with the exact mean $\mathbb{E}\left[\hat{z}^{\varepsilon}_{j,k}\right]$. The mother wavelet is a Daubechies wavelet with 5 vanishing moments. Moreover, the number of observations, N, is 129.

That example of the tent map is a good illustration of the complexity of a chaotic signal with measurement noise: its filtering is more complex than for a signal with a linear noise influence. Therefore, a signal with non-linear noise influence has not to be filtered with the same methods.

3.3 The Hénon map

In the following example, we propose a chaos in higher dimension, the Hénon map, which is described by:

$$\begin{cases} x_{n+1} &= y_n + 1 - ax_n^2 \\ y_{n+1} &= bx_n. \end{cases}$$

If $a = 1.4$ and $b = 0.3$, then $(x_n, y_n)_{n\in\mathbb{N}}$ is a chaotic system [12], whose phase plot is represented in the Figure 5.

We consider a Gaussian noise and we represent the results in the Figure 6. We have chosen the two-dimensional Daubechies mother wavelet with 5 vanishing moments and the wavelet coefficient $j = 3$, $k = (4, 4)$ and $\eta = 2$: $\hat{z}^{\varepsilon}_{2,3,(4,4)}$. For a centred Gaussian noise with standard deviation $\sigma = 5\%$ and a sample size $N = 200$, we obtain the following wavelet coefficient for the pure signal[12]: $\hat{z}_{2,3,(4,4)} = 2.00 \times 10^{-3}$. We can observe the impact of the noise on $\hat{z}_{2,3,(4,4)}$ in the Figure 6.

For the dynamic or the measurement noise environment, we have a good fit between the empirical distribution of the wavelet coefficient $\hat{z}^{\varepsilon}_{2,3,(4,4)}$ and its exact distribution obtained by Theorem 4 and Theorem 5. Thus the multidimensional case seems well handled by the approximations we proposed.

[12] This empirical value is obtained by taking arbitrarily the same Voronoi region size for all the observations.

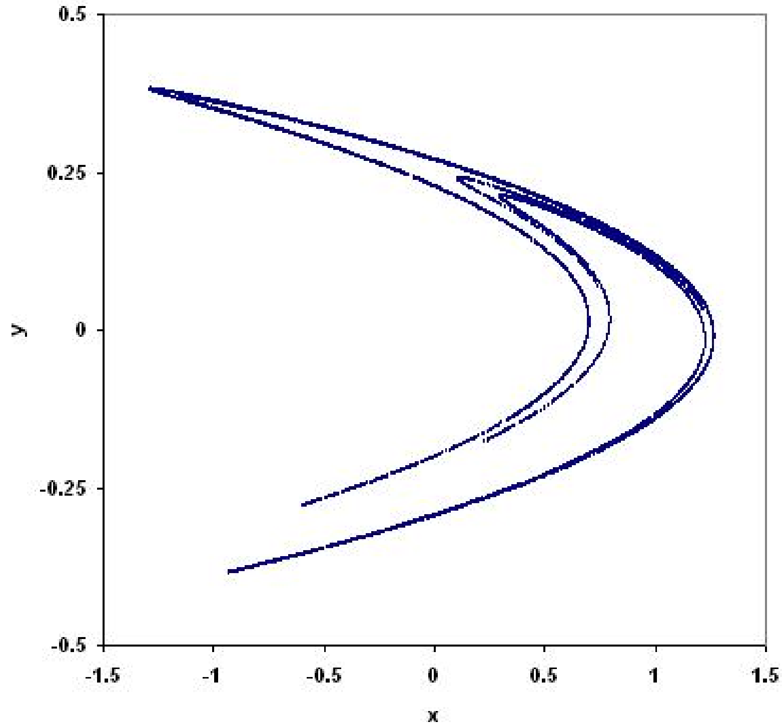

Figure 5: Phase plot of the chaotic Hénon map with $a = 1.4$ and $b = 0.3$.

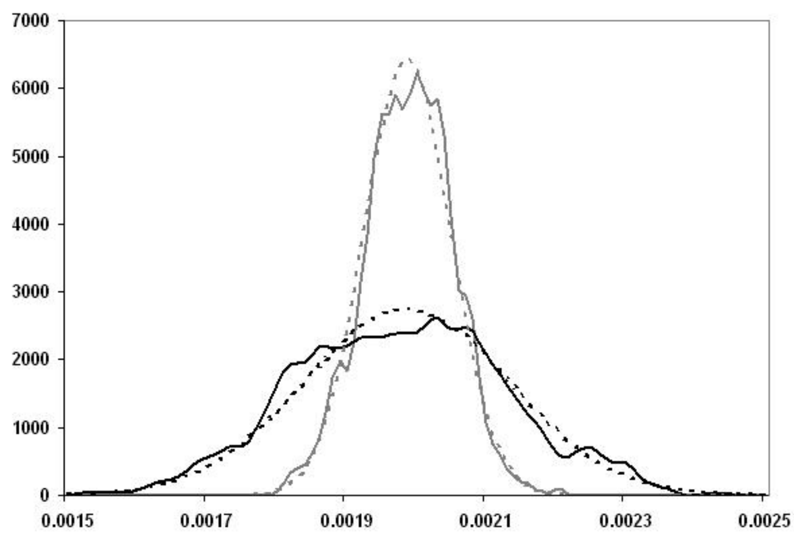

Figure 6: In grey, probability density function of the wavelet coefficient $\hat{z}^{\varepsilon}_{2,3,(4,4)}$ of the Hénon map of parameters $a = 1.4$ and $b = 0.3$ for a centred Gaussian dynamic noise with standard deviation 0.05. If the dynamic noise is replaced by a centred Gaussian measurement noise with standard deviation 0.05, we get the black lines. In both cases, the solid thick line is the empirical density obtained by $1,000$ simulations and represented with a Gaussian kernel, the dotted line is the exact density obtained by Theorem 4 for the dynamic noise and the approximation obtained by Proposition 5 for the measurement noise. The mother wavelet is a Daubechies wavelet with 5 vanishing moments. Moreover, the number of observations, N, is 200.

4 Conclusion

We presented several methods to get the probability distribution of each wavelet coefficients of a noisy chaos. If we face a dynamical noise, Theorem 1 provide an exact expression. For measurement noise, two approximations are proposed in Theorem 2 and Algorithms 1 and 2. Those approximations are of course more accurate when the noise is small and we are able to quantify what *small* means: the variance of the wavelet coefficient, which depends on the chaos, must be lower than 2. For au Gaussian noise, for the mother wavelet considered and the resolution and translation parameters considered in the article, this corresponds approximately to a variance of the noise equal to 48 for a tent map and equal to 155 for a logistic map.

Moreover, using the multidimensional wavelets, we generalize these results to multidimensional chaos in Theorem 4 and Theorem 5. This allows us to present a quite wide set of examples, from the logistic map to the Hénon map.

These results are very useful to determine the optimal threshold to use when filtering a noisy chaotic attractor to recover the pure chaotic attractor. This is the aim of another article [9].

References

[1] Barndorff-Nielsen, O. (1977), *Exponentially decreasing distributions for the logarithm of particle size*, Proceedings of the Royal Society London A, 353: 401-419

[2] Barndorff-Nielsen, O. and Halgreen, C. (1977), *Infinite divisibility of the hyperbolic and generalized inverse Gaussian distribution*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 38: 309-312

[3] Blacher, R. (2003), *Multivariate quadratic forms of random vectors*, Journal of Multivariate Analysis, 87, 1: 2-23

[4] Breeden, J., Dinkelacker, F. and Hübler, A. (1990), *Noise in the modeling and control of dynamical systems*, Physical review A, 42: 5827-5836

[5] Cornish, E. and Fisher, R. (1937), *Moments and cumulants in the specification of distributions*, Revue de l'Institut International de Statistique, 5: 307-320

[6] Daubechies, I. (1992), *Ten lectures on wavelets*, SIAM, Philadelphia

[7] Donoho, D. and Johnstone, I. (1994), *Ideal spatial adaptation via wavelet shrinkage*, Biometrika, 81: 425-455

[8] Garcin, M. and Guégan, D. (2012), *Extreme values of random or chaotic discretization steps and connected networks*, Applied Mathematical Sciences, 119, 6: 5901-5926

[9] Garcin, M. and Guégan, D. (2013), *Optimal wavelet shrinkage of a noisy chaos*, to appear

[10] Georgiou, P., Tskalides, P. and Kyriakakis, C. (1999), *Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise*, IEEE transactions on multimedia, 1, 3: 291-301

[11] Gil, A., Segura, J. and Temme, N. (2007), *Numerical methods for special functions*, SIAM, Philadelphia

[12] Guégan, D. (2003), *Les Chaos en finance : approche statistique*, Economica, Paris

[13] Guégan, D. (2008), *Effect of noise filtering on predictions: on the routes of chaos*, CES working paper #8

[14] Guégan, D. and Hoummiya, K. (2005), *Denoising with wavelets method in chaotic time series: application in climatology, energy and finance* in *Noise and fluctuations in econophysics and finance*, ed. by D. Abbott, J.P. Bouchaud, X. Gabaix, J.L. McCaulay, Proc. SPIE 5848: 174-185

[15] de Haan, L. and Ferreira, A. (2006), *Extreme value theory: an introduction*, Springer, New York

[16] Ilow, J. (1995), *Signal processing in alpha-stable noise environments: noise modeling, detection and estimation*, Ph.D. dissertation, Univ. Toronto, Toronto

[17] Ilow, J. and Hatzinakos, D. (1998), *Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers*, IEEE transactions on signal processing, 46, 6: 1601-1611

[18] Kalluri, S. and Arce, G. (1998), *Adaptative weighted median filter algorithms for robust signal processing in alpha-stable noise environments*, IEEE transactions on signal processing, 46: 322-334

[19] Lacoume, J.-L., Amblard, P.-O. and Comon, P. (1997), *Statistiques d'ordres supérieurs pour le traitement du signal*, Masson, Paris

[20] Mallat, S. (2000), *Une exploration des signaux en ondelettes*, Ellipses, Éditions de l'École Polytechnique, Paris

[21] Mason, J. and Handscomb, D. (2003), *Chebyshev polynomials*, Chapman & Hall/CRC, New York

[22] Rioul, O. and Vetterli, M. (1991), *Wavelets and signal processing*, IEEE signal processing magazine, 8: 14-38

[23] Samorodnitsky, G. and Taqqu, M. (1994), *Stable non-Gaussian random Processes*, Chapman & Hall, New York

[24] Schoukens, J. and Renneboog, J. (1986), *Modeling the noise influence on the Fourier coefficients after a discrete Fourier transform*, IEEE transactions on instrumentation and measurement, 25, 3: 278-286

[25] Trefethen, L. (2013), *Approximation theory and approximation practice*, SIAM, Philadelphia

A Proof of Theorem 1

Proof. The wavelet coefficient $\hat{z}^{\varepsilon}_{j,k}$ of z^{ε} is:

$$\begin{array}{rcl}\hat{z}^{\varepsilon}_{j,k} & = & \sum_{n=1}^{N} z^{\varepsilon}(u_n)\psi_{j,k}(u_n)(u_n-u_{n-1}) \\ & = & \sum_{n=1}^{N} \left[z(u_n)+\varepsilon_n\right]\psi_{j,k}(u_n)(u_n-u_{n-1}) \\ & = & \hat{z}_{j,k} + \sum_{n=1}^{N} \varepsilon_n\psi_{j,k}(u_n)(u_n-u_{n-1}).\end{array}$$

Let $t \in \mathbb{R}$. The characteristic function of $\hat{z}^{\varepsilon}_{j,k}$ is then:

$$\begin{array}{rcl}\mathbb{E}\left[e^{it\hat{z}^{\varepsilon}_{j,k}}\right] & = & \exp\left(it\hat{z}_{j,k}\right)\mathbb{E}\left[\prod_{n=1}^{N}\exp\left(it\varepsilon_n\psi_{j,k}(u_n)(u_n-u_{n-1})\right)\right] \\ & = & \exp\left(it\hat{z}_{j,k}\right)\prod_{n=1}^{N}\mathbb{E}\left[\exp\left(it\varepsilon_n\psi_{j,k}(u_n)(u_n-u_{n-1})\right)\right],\end{array}$$

since the variables $\varepsilon_1, \varepsilon_2, ..., \varepsilon_N$ are independent. Then, using the equation (4), we get:

$$\begin{array}{rcl}\mathbb{E}\left[e^{it\hat{z}^{\varepsilon}_{j,k}}\right] & = & \exp\left(it\hat{z}_{j,k}\right)\prod_{n=1}^{N}\exp\left(i\mu t\psi_{j,k}(u_n)(u_n-u_{n-1}) - \gamma|t\psi_{j,k}(u_n)(u_n-u_{n-1})|^{\alpha}\right) \\ & = & \exp\left(it\hat{z}_{j,k} + \sum_{n=1}^{N} i\mu t\psi_{j,k}(u_n)(u_n-u_{n-1}) - \sum_{n=1}^{N}\gamma|t\psi_{j,k}(u_n)(u_n-u_{n-1})|^{\alpha}\right) \\ & = & \exp\left(i\mu' t - \gamma'|t|^{\alpha}\right),\end{array}$$

where:

$$\left\{ \begin{array}{lcl} \gamma' & = & \gamma \sum_{n=1}^{N} |\psi_{j,k}(u_n)(u_n - u_{n-1})|^{\alpha} \\ \mu' & = & \hat{z}_{j,k} + \mu \sum_{n=1}^{N} \psi_{j,k}(u_n)(u_n - u_{n-1}). \end{array} \right.$$

We recognize the characteristic function of a symmetric alpha-stable random variable of parameters α, γ', μ', like in equation (4). □

B Proof of Proposition 1

Proof. This is a direct consequence of Theorem 1 since any centred Gaussian random variable of variance σ^2 is also a symmetric 2-stable random variable with parameters $\mu = 0$ and $\gamma = \sigma^2/2$. We also note that, when $N \to \infty$, the maximum discretization step size is decreasing towards 0 because of the ergodic condition on the chaos whose attractor is z. A more precise expression of that maximum step size may be found in [8]. Nevertheless, when $N \to \infty$:

$$\begin{array}{lcl} \sigma' & \overset{N\to\infty}{\longrightarrow} & \lim\limits_{N\to\infty} \sigma \sqrt{\sum_{n=1}^{N} \left| \int_{u_{n-1}}^{u_n} \psi_{j,k}(u) du \right|^2} \\ & \leq & \lim\limits_{N\to\infty} \sigma \sqrt{\sum_{n=1}^{N} \int_{u_{n-1}}^{u_n} |\psi_{j,k}(u)|^2 \, du} \end{array}$$

thanks to Jensen's inequality. Then:

$$\begin{array}{lcl} \sigma' & \leq & \lim\limits_{N\to\infty} \sigma \sqrt{\int_{u_0}^{u_N} |\psi_{j,k}(u)|^2 \, du} \\ & \leq & \sigma \sqrt{\int_{\mathbb{R}} |\psi_{j,k}(u)|^2 \, du} = \sigma \|\psi_{j,k}\|_{\mathcal{L}^2(\mathbb{R})}. \end{array}$$

Finally, if $\psi_{j,k}$ is an orthonormal wavelet we simply get $\sigma' \leq \sigma$. □

C Proof of Theorem 2

Proof. From the equation (9), we can achieve a first-order Taylor expansion on z, which is continuous and is continuously differentiable on each segment included in $\mathbb{R} \setminus \mathcal{N}$ and which is therefore continuous and continuously differentiable in each u_n, because they are assumed to be in $\mathbb{R} \setminus \mathcal{N}$:

$$\forall n \in \{1, ..., N\}, \; z^{\varepsilon}(u_n) \overset{\varepsilon_n^{\star}\to 0}{\sim} z(u_n) - z'(u_n)\varepsilon_n^{\star} + \varepsilon_n. \tag{12}$$

In the equation (12), we used the notation $a(x) \overset{x\to 0}{\sim} b(x)$, for functions a and b. This means that, when $\lim\limits_{x\to 0} b(x) \neq 0$, then $\lim\limits_{x\to 0} \frac{a(x)}{b(x)} = 1$. Using the equation (12), the wavelet coefficient $\hat{z}^{\varepsilon}_{j,k}$ of z^{ε} is:

$$\begin{array}{lcl} \hat{z}^{\varepsilon}_{j,k} & = & \sum_{n=1}^{N} z^{\varepsilon}(u_n) \psi_{j,k}(u_n)(u_n - u_{n-1}) \\ & \overset{\max\limits_{n\in\{1,\ldots,N\}} |\varepsilon_n^{\star}|\to 0}{\sim} & \sum_{n=1}^{N} \left[z(u_n) - z'(u_n)\varepsilon_n^{\star} + \varepsilon_n \right] \psi_{j,k}(u_n)(u_n - u_{n-1}) \\ & \overset{\max\limits_{n\in\{1,\ldots,N\}} |\varepsilon_n^{\star}|\to 0}{\sim} & \hat{z}_{j,k} + \sum_{n=1}^{N} \left[-z'(u_n)\varepsilon_n^{\star} + \varepsilon_n \right] \psi_{j,k}(u_n)(u_n - u_{n-1}). \end{array}$$

Then, the limit of the random variable $\hat{z}^{\varepsilon}_{j,k}$ when $\max\limits_{n\in\{1,\ldots,N\}} |\varepsilon_n^{\star}| \to 0$, denoted $\hat{Z}^{\varepsilon}_{j,k}$, is simply:

$$\hat{Z}^{\varepsilon}_{j,k} = \hat{z}_{j,k} + \sum_{n=1}^{N} \left[-z'(u_n)\varepsilon_n^{\star} + \varepsilon_n \right] \psi_{j,k}(u_n)(u_n - u_{n-1}). \tag{13}$$

Let $t \in \mathbb{R}$. The characteristic function of $\hat{Z}^{\varepsilon}_{j,k}$ is then:

$$\begin{array}{lcl} \mathbb{E}\left[e^{it\hat{Z}^{\varepsilon}_{j,k}} \right] & = & \exp\left(it\hat{z}_{j,k}\right) \mathbb{E}\left[\prod_{n=1}^{N} \exp\left(it\varepsilon_n \psi_{j,k}(u_n)(u_n - u_{n-1})\right) \exp\left(-itz'(u_n)\varepsilon_n^{\star} \psi_{j,k}(u_n)(u_n - u_{n-1})\right) \right] \\ & = & \exp\left(it\hat{z}_{j,k}\right) \prod_{n=1}^{N} \mathbb{E}\left[\exp\left(it\varepsilon_n \psi_{j,k}(u_n)(u_n - u_{n-1})\right) \right] \mathbb{E}\left[\exp\left(-itz'(u_n)\varepsilon_n^{\star} \psi_{j,k}(u_n)(u_n - u_{n-1})\right) \right], \end{array} \tag{14}$$

since the variables $\varepsilon_1, \varepsilon_2, ..., \varepsilon_N, \varepsilon_1^\star, \varepsilon_2^\star, ..., \varepsilon_N^\star$ are independent. Then, using the equation (4) in the equation (14), we get:

$$\begin{array}{rcl}\mathbb{E}\left[e^{it\hat{Z}^\varepsilon_{j,k}}\right] & = & \exp\left(it\hat{z}_{j,k}\right)\prod_{n=1}^N \exp\left(i\mu t\psi_{j,k}(u_n)(u_n-u_{n-1}) - \gamma|t\psi_{j,k}(u_n)(u_n-u_{n-1})|^\alpha\right) \\ & & \prod_{n=1}^N \exp\left(-i\mu t z'(u_n)\psi_{j,k}(u_n)(u_n-u_{n-1}) - \gamma|z'(u_n)|^\alpha|t\psi_{j,k}(u_n)(u_n-u_{n-1})|^\alpha\right) \\ & = & \exp\left(it\hat{z}_{j,k} + \sum_{n=1}^N i\mu t(1-z'(u_n))\psi_{j,k}(u_n)(u_n-u_{n-1}) - \sum_{n=1}^N \gamma(1+|z'(u_n)|^\alpha)|t\psi_{j,k}(u_n)(u_n-u_{n-1})|^\alpha\right) \\ & = & \exp\left(i\mu' t - \gamma'|t|^\alpha\right),\end{array}$$

where:

$$\begin{cases} \gamma' & = & \gamma\sum_{n=1}^N \left(1+|z'(u_n)|^\alpha\right)|\psi_{j,k}(u_n)(u_n-u_{n-1})|^\alpha \\ \mu' & = & \hat{z}_{j,k} + \mu\sum_{n=1}^N \left(1-z'(u_n)\right)\psi_{j,k}(u_n)(u_n-u_{n-1}). \end{cases}$$

We recognize the characteristic function of a symmetric alpha-stable random variable of parameters α, γ', μ', like in equation (4). □

D Proof of Theorem 3

Proof. Let r be the remainder of the Taylor-Lagrange expansion of z^ε:

$$\forall n \in \{1, ..., N\},\ z(u_n - \varepsilon_n^\star) = z(u_n) - z'(u_n)\varepsilon_n^\star + r(n),$$

where

$$|r(n)| \leq \frac{|\varepsilon_n^\star|^2 \max\limits_{x\in Supp(z)}\left(|z''(x)|\right)}{2}.$$

Then,

$$z^\varepsilon(u_n) = z(u_n) - z'(u_n)\varepsilon_n^\star + \varepsilon_n + r(n)$$

and we get:

$$\begin{array}{rcl} d_{j,k} & = & \left|\hat{z}^\varepsilon_{j,k} - \hat{Z}^\varepsilon_{j,k}\right| \\ & = & \left|\sum_{n=1}^N r(n)\psi_{j,k}(u_n)(u_n-u_{n-1})\right| \\ & \leq & \frac{1}{2}\max\limits_{x\in Supp(z)}\left(|z''(x)|\right)\sum_{n=1}^N |\varepsilon_n^\star|^2\,|\psi_{j,k}(u_n)(u_n-u_{n-1})|. \end{array} \tag{15}$$

From the equation (15), two possible bounds arise. The first one leads to an approximation of the distribution of the error when N tends towards infinity. The second one works particularly well when the number of observations is limited.

1. We can keep the wavelet function in the sum in the equation (15) and extract the noise from that sum:

$$d_{j,k} \leq \frac{1}{2}\max_{x\in Supp(z)}\left(\left|z''(x)\right|\right)\max_{n\in\{1,...,N\}}\left(\left|\varepsilon_n^\star\right|^2\right)\sum_{n=1}^N \left|\psi_{j,k}(u_n)\right|(u_n-u_{n-1}). \tag{16}$$

   For an ergodic chaos, when N is big, the biggest step size $(u_n - u_{n-1})$ is small, so that we can write:

$$\sum_{n=1}^N \left|\psi_{j,k}(u_n)\right|(u_n-u_{n-1}) \overset{N\to\infty}{\longrightarrow} \int_{u_0}^{u_N}\left|\psi_{j,k}(x)\right|dx \leq \int_{\mathbb{R}}\left|\psi_{j,k}(x)\right|dx. \tag{17}$$

   Then, using together equation (16) and equation (17), we get: $\forall\epsilon > 0,\ \exists N' \in \mathbb{N}$ such that when $N \geq N'$

$$d_{j,k} \leq \frac{1}{2}\max_{x\in Supp(z)}\left(\left|z''(x)\right|\right)M_{|\varepsilon^\star|^2}\int_{\mathbb{R}}\left|\psi_{j,k}(x)\right|dx + \epsilon, \tag{18}$$

   where

$$M_{|\varepsilon^\star|^2} = \lim_{N\to\infty}\left(\max_{n\in\{1,...,N\}}\left(\left|\varepsilon_n^\star\right|^2\right)\right)$$

   is a random variable whose distribution may be explicitly known thanks to the extreme value theory.

2. On the contrary, we can keep the noise in the sum of the equation (15) and extract the wavelet function from that sum:

$$d_{j,k} \leq \frac{1}{2} \max_{x \in Supp(z)} \left(|z''(x)|\right) \max_{x \in Supp(\psi_{j,k})} \left(|\psi_{j,k}(x)|\right) \max_{n \in \{1,\dots,N\}} (u_n - u_{n-1}) S_{N,|\varepsilon^\star|^2}, \quad (19)$$

where

$$S_{N,|\varepsilon^\star|^2} = \sum_{n=1}^{N} |\varepsilon_n^\star|^2.$$

In equation (19), depending on the cumulative distribution function of the observations, Δ_U, we can find the cumulative distribution function of $\max\limits_{n \in \{1,\dots,N\}} (u_n - u_{n-1})$. In particular, even when N does not tends towards infinity, we can use the bound presented in [8]:

$$\Delta_{\max\limits_{n \in \{1,\dots,N\}} (u_n - u_{n-1})} \geq \max \left\{ 0, 1 - N \int_{\inf(Supp(\delta_U))+v}^{\sup(Supp(\delta_U))} \delta_U(u) \left[\Delta_U(u-v) + 1 - \Delta_U(u)\right]^{N-1} du \right\}, \quad (20)$$

where δ_U is the probability density function corresponding to the cumulative distribution function Δ_U. In addition to that, we note that if X and Y are positive independent random variables, and $x, y > 0$, then

$$X \leq x \text{ and } Y \leq y \Rightarrow XY \leq xy$$

which leads to:

$$\mathbb{P}(XY \leq xy) \geq \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y). \quad (21)$$

In particular, in equation (19), using equation (21), we have:

$$\Delta_{\max\limits_{n \in \{1,\dots,N\}} (u_n - u_{n-1}) S_{N,|\varepsilon^\star|^2}}(xy) \geq \Delta_{\max\limits_{n \in \{1,\dots,N\}} (u_n - u_{n-1})}(x) \Delta_{S_{N,|\varepsilon^\star|^2}}(y). \quad (22)$$

On one hand thanks to equation (18) where ϵ can be arbitrarily small and on the other hand thanks to equations (19), (20) and (22), we can prove the result of Theorem 3, namely:

$$\begin{cases} D_{j,k}(p) & \leq & \Omega_N^\infty(p) \\ D_{j,k}(p) & \leq & \Omega_N(p). \end{cases}$$

□

E Proof of Proposition 5

Proof. g_N is the cumulative distribution function of $\sum_{n=1}^{N} |\varepsilon_n^\star|^2$ and h_N is the cumulative distribution function of $\lim\limits_{N \to \infty} \left(\max\limits_{n \in \{1,\dots,N\}} \left(|\varepsilon_n^\star|^2\right) \right)$.

1. $\sum_{n=1}^{N} \left| \frac{\varepsilon_n^\star}{\sigma} \right|^2$ is a chi-squared random variable with N degrees of freedom. Thus:

$$g_N : x \geq 0 \mapsto \frac{\gamma\left(\frac{N}{2}, \frac{\sigma^2 x}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}.$$

2. For a random variable X, we denote Δ_X its cumulative distribution function and δ_X the probability density function. Since $\varepsilon_1^\star$ is Gaussian, then $|\varepsilon_1^\star/\sigma|^2$ is a chi-squared random variable with only one degree of freedom. We denote Y a random variable with the same probability density function than $|\varepsilon_1^\star|^2$. Let $x > 0$. Then:

$$\delta_Y(x) = \frac{1}{\sigma\sqrt{2\pi x}} \exp\left(-\frac{x}{2\sigma^2}\right). \quad (23)$$

In order to apply some theorem from the extreme value theory, more particularly the von Mises' condition [15], we need to find the limit, when $x \to \infty$, of the derivative of the function D, defined by:

$$D(x) = \frac{1 - \Delta_Y(x)}{\delta_Y(x)}. \tag{24}$$

Its derivative is:

$$D'(x) = \frac{-\delta_Y(x)^2 - (1 - \Delta_Y(x))\delta'_Y(x)}{\delta_Y(x)^2}. \tag{25}$$

From the equation (23), we calculate that:

$$\begin{array}{lcl} \delta'_Y(x) & = & -\frac{1}{2\sigma\sqrt{2\pi x}}\left(\frac{1}{x} + \frac{1}{\sigma^2}\right)\exp\left(-\frac{x}{2\sigma^2}\right) \\ & \overset{x\to\infty}{\sim} & -\frac{1}{2\sigma^3\sqrt{2\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right). \end{array} \tag{26}$$

We now have to find an equivalent, when $x \to \infty$, of $1 - \Delta_Y(x)$. First, let $y > x$ and I the function defined by:

$$I(x, y) = \int_x^y \delta_Y(s) ds. \tag{27}$$

Integrating the equation (27) by parts, we get:

$$I(x, y) = \sigma\sqrt{\frac{2}{\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right) - \sigma\sqrt{\frac{2}{\pi y}}\exp\left(-\frac{y}{2\sigma^2}\right) - \int_x^y \frac{2\sigma^2}{2s}\frac{1}{\sigma\sqrt{2\pi s}}\exp\left(-\frac{s}{2\sigma^2}\right) ds, \tag{28}$$

so that, making $y \to \infty$ in the equation (28):

$$1 - \Delta_Y(x) = \sigma\sqrt{\frac{2}{\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right) - \int_x^\infty \frac{\sigma^2}{s}\frac{1}{\sigma\sqrt{2\pi s}}\exp\left(-\frac{s}{2\sigma^2}\right) ds. \tag{29}$$

Since $x, \sigma > 0$, we note that:

$$\left|\int_x^\infty \frac{\sigma^2}{s}\frac{1}{\sigma\sqrt{2\pi s}}\exp\left(-\frac{s}{2\sigma^2}\right) ds\right| \leq \frac{\sigma^2}{x}(1 - \Delta_Y(x)), \tag{30}$$

so that, using the equations (29) and (30), we get the equivalence relation:

$$1 - \Delta_Y(x) \overset{x\to\infty}{\sim} \sigma\sqrt{\frac{2}{\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right). \tag{31}$$

Finally, from the equations (23), (25), (26) and (31) we get:

$$D'(x) \overset{x\to\infty}{\sim} -1 - \frac{\left[\sigma\sqrt{\frac{2}{\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right)\right]\left[-\frac{1}{2\sigma^3\sqrt{2\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right)\right]}{\left[\frac{1}{\sigma\sqrt{2\pi x}}\exp\left(-\frac{x}{2\sigma^2}\right)\right]^2} = 0. \tag{32}$$

Then, using the equation (32) and von Mises' condition, we can assert that Δ_Y is in the max-domain of attraction of a GEV of parameter 0, namely a Gumbel distribution: there exists a real normalization series (a_N, b_N) such that the distribution function of the random variable $b_N^{-1}\left(\max_{n\in\{1,\ldots,N\}}\left(|\varepsilon_n^\star|^2\right) - a_N\right)$ converges towards the distribution function of an affine transformation of a Gumbel random variable, whose distribution function is:

$$G_0 : x \longrightarrow \exp\left(-e^{-x}\right). \tag{33}$$

The parameters of the affine transformation will be determined accurately next in this proof. Moreover, von Mises' condition allows us to write that:

$$\left\{\begin{array}{lclcl} a_N & \overset{N\to\infty}{\sim} & \Delta_Y^{-1}\left(1 - \frac{1}{N}\right) & \overset{N\to\infty}{\longrightarrow} & +\infty \\ b_N & \overset{N\to\infty}{\sim} & D(a_N) & \overset{N\to\infty}{\longrightarrow} & 2\sigma^2, \end{array}\right. \tag{34}$$

thanks to the equations (23), (24) and (31). Let α be a function defined by:

$$\alpha(N)^{\frac{1}{2}}e^{\alpha(N)} = \frac{N}{\sqrt{\pi}}.$$

One can numerically easily get α but also a proof of its existence and uniqueness, the function $t \mapsto t^{\frac{1}{2}}e^t$ being a bijection of $[0,\infty)$. Then, thanks to the equation (31), we remark that:

$$1 - \Delta_Y\left(2\sigma^2\alpha(N)\right) \overset{N\to\infty}{\sim} \frac{1}{N},$$

so that, Δ_Y being an increasing function, we can choose to define a_N by:

$$a_N = 2\sigma^2\alpha(N). \tag{35}$$

The limit of the translated and scaled maximum of the set of random variables can now be written. Since $b_N x = \overset{N\to\infty}{o}(a_N)$, then:

$$\begin{array}{rcl} \mathbb{P}\left(\frac{\max\limits_{n\in\{1,\dots,N\}}\left(|\varepsilon_n^\star|^2\right)-a_N}{b_N} \leq x\right) & = & \left[\Delta_Y(xb_N + a_N)\right]^N \\ & \overset{N\to\infty}{\sim} & \left[1-(1-\Delta_Y(a_N))\exp\left(-\frac{xb_N}{2\sigma^2}\right)\right]^N \\ & \overset{N\to\infty}{\sim} & \left[1-\frac{1}{N}\exp\left(-x\right)\right]^N \\ & \overset{N\to\infty}{\longrightarrow} & G_0(x). \end{array} \tag{36}$$

Finally, using the equations (33), (34), (35) and (36), we get:

$$\begin{array}{rcl} h_N(x) & \overset{N\to\infty}{\sim} & G_0\left(\frac{x-a_N}{b_N}\right) \\ & \overset{N\to\infty}{\sim} & \exp\left(-\exp\left(\alpha(N)-\frac{x}{2\sigma^2}\right)\right). \end{array}$$

□

We remark that successive Taylor expansions using the equation (31) can lead to build α as the limit of the sequence of functions $(\alpha_k)_{k\in\mathbb{N}}$, defined by:

$$\left\{\begin{array}{rcl} \alpha_0(N) & = & \log\left(\frac{N}{\sqrt{\pi}}\right) \\ \alpha_{k+1}(N) & = & \log\left(\frac{N}{\sqrt{\pi\alpha_k(N)}}\right). \end{array}\right. \tag{37}$$

Then α is the limit of α_k, when $k\to\infty$. Indeed, when we replace α_{k+1} and α_k by α in the equation (37), we find that α is such that:

$$\alpha(N)^{\frac{1}{2}}e^{\alpha(N)} = \frac{N}{\sqrt{\pi}},$$

what is consistent with its definition in Proposition 5.

F Proof of Proposition 6

Proof. Let $x \geq 0$. We denote Y a random variable with the same probability density function than $|\varepsilon_1^\star|^2$. We remark that:

$$\begin{array}{rcl} \Delta_Y(x) & = & \mathbb{P}\left(|\varepsilon_1^\star| \leq \sqrt{x}\right) \\ & = & \Delta_{\varepsilon_1^\star}(\sqrt{x}) - \Delta_{\varepsilon_1^\star}(-\sqrt{x}) \\ & = & \frac{2}{\pi}\arctan\left(\frac{\sqrt{x}}{\gamma}\right), \end{array}$$

because the cumulative distribution function of a Cauchy random variable with scale parameter γ is:

$$\Delta_{\varepsilon_1^\star}(x) = \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x}{\gamma}\right).$$

Therefore, the density probability function of Y is:

$$\begin{array}{rcl} \delta_Y(x) & = & \frac{\partial\Delta_{\varepsilon_1^\star}(\sqrt{x})}{\partial x} - \frac{\partial\Delta_{\varepsilon_1^\star}(-\sqrt{x})}{\partial x} \\ & = & \frac{\partial}{\partial x}\left[\frac{2}{\pi}\arctan\left(\frac{\sqrt{x}}{\gamma}\right)\right] \\ & = & \frac{1}{\gamma\pi\sqrt{x}\left(1+\frac{x}{\gamma^2}\right)} \\ & \overset{x\to\infty}{\sim} & \frac{\gamma}{\pi x^{3/2}}, \end{array} \tag{38}$$

Like for Proposition 5, we are going to use the von Mises' condition: we look for the limit, when $x \to \infty$, of the derivative of the function D, defined by:

$$D(x) = \frac{1-\Delta_Y(x)}{\delta_Y(x)}.$$

Its derivative is:

$$D'(x) = \frac{-\delta_Y(x)^2 - (1-\Delta_Y(x))\delta_Y'(x)}{\delta_Y(x)^2}. \tag{39}$$

From the equation (38), we calculate that:

$$\begin{array}{rcl} \delta_Y'(x) & = & -\frac{3x+\gamma^2}{2\pi\gamma^3 x^{3/2}\left(1+\frac{x}{\gamma^2}\right)^2} \\ & \overset{x\to\infty}{\sim} & -\frac{3\gamma}{2\pi x^{5/2}}. \end{array} \tag{40}$$

Moreover, integrating the equation (38), we have an other approximation:

$$\begin{array}{rcl} 1-\Delta_Y(x) & \overset{x\to\infty}{\sim} & \int_x^\infty \frac{\gamma}{\pi y^{3/2}} dy \\ & \overset{x\to\infty}{\sim} & \frac{2\gamma}{\pi\sqrt{x}}. \end{array} \tag{41}$$

From the equations (38), (39), (40) and (41) we get an equivalent of the derivative of D:

$$D'(x) \overset{x\to\infty}{\sim} \frac{-\left[\frac{\gamma}{\pi x^{3/2}}\right]^2 - \left[\frac{2\gamma}{\pi\sqrt{x}}\right]\left[-\frac{3\gamma}{2\pi x^{5/2}}\right]}{\left[\frac{\gamma}{\pi x^{3/2}}\right]^2} = 2.$$

According to von Mises' condition, we can assert that Δ_Y is in the max-domain of attraction of a GEV of parameter 2, namely a Fréchet distribution: there exists a real normalization series (a_N, b_N) such that the distribution function of the random variable $b_N^{-1}\left(\max_{n\in\{1,\dots,N\}}\left(|\varepsilon_n^\star|^2\right) - a_N\right)$ converges towards the distribution function of an affine transformation of a Fréchet random variable of parameter $1/2$, whose distribution function is:

$$G_{1/2} : x \longrightarrow \exp\left(-x^{-1/2}\right). \tag{42}$$

The parameters of the affine transformation will be determined accurately next in this proof. Moreover, von Mises' condition allows us to write that:

$$\left\{ \begin{array}{lcl} a_N & \overset{N\to\infty}{\sim} & \Delta_Y^{-1}\left(1-\frac{1}{N}\right) \\ b_N & \overset{N\to\infty}{\sim} & D(a_N). \end{array} \right. \tag{43}$$

The equation (41) incites us to choose:

$$a_N = \frac{4\gamma^2 N^2}{\pi^2}, \tag{44}$$

from which we get, using the equations (38) and (43):

$$b_N = \frac{8\gamma^2 N^2}{\pi^2}. \tag{45}$$

The limit of the translated and scaled maximum of the set of random variables can now be written, using the equations (41), (44) and (45):

$$\begin{array}{rcl}\mathbb{P}\left(\frac{\max\limits_{n\in\{1,...,N\}}\left(\left|\varepsilon_n^\star\right|^2\right)-a_N}{b_N}\leq x\right) & = & \left[\Delta_Y(xb_N+a_N)\right]^N \\ & \overset{N\to\infty}{\sim} & \left[1-\frac{2\gamma}{\pi\sqrt{b_N x+a_N}}\right]^N \\ & \overset{N\to\infty}{\sim} & \left[1-\frac{(2x+1)^{-1/2}}{N}\right]^N \\ & \overset{N\to\infty}{\longrightarrow} & G_{1/2}(2x+1).\end{array} \tag{46}$$

Finally, using the equations (42), (44), (45) and (46), we get:

$$\begin{array}{rcl}h_N(x) & \overset{N\to\infty}{\sim} & G_{1/2}\left(2\left[\frac{x-a_N}{b_N}\right]+1\right) \\ & \overset{N\to\infty}{\sim} & \exp\left(-\frac{2\gamma N}{\pi\sqrt{x}}\right).\end{array}$$

□

G Proof of Theorem 4

Proof. The wavelet coefficient $\hat{Z}^\varepsilon_{\eta,j,k}$ of Z^ε is:

$$\begin{array}{rcl}\hat{Z}^\varepsilon_{\eta,j,k} & = & \sum_{n=1}^N Z^\varepsilon(U_n)\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx \\ & = & \sum_{n=1}^N\left[Z(U_n)+(\varepsilon_{1,n},...,\varepsilon_{p,n})\right]\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx \\ & = & \hat{Z}_{\eta,j,k}+\left(\sum_{n=1}^N\varepsilon_{1,n}\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx,...,\sum_{n=1}^N\varepsilon_{1,n}\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx\right).\end{array}$$

Let $t\in\mathbb{R}$. The characteristic function of the m-th coordinate of $\hat{Z}^\varepsilon_{\eta,j,k}$ is then:

$$\begin{array}{rcl}\mathbb{E}\left[\exp\left(it\left(\hat{Z}^\varepsilon_{\eta,j,k}\right)_m\right)\right] & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m\right)\mathbb{E}\left[\prod_{n=1}^N\exp\left(it\varepsilon_{m,n}\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx\right)\right] \\ & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m\right)\prod_{n=1}^N\mathbb{E}\left[\exp\left(it\varepsilon_{m,n}\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx\right)\right],\end{array}$$

since the variables $\varepsilon_{m,1},\varepsilon_{m,2},...,\varepsilon_{m,N}$ are independent. Then, using the equation (4), we get:

$$\begin{array}{rcl}\mathbb{E}\left[\exp\left(it\left(\hat{Z}^\varepsilon_{\eta,j,k}\right)_m\right)\right] & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m\right)\prod_{n=1}^N\exp\left(i\mu t\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx-\gamma|t\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx|^\alpha\right) \\ & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m+\sum_{n=1}^N i\mu t\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx-\sum_{n=1}^N\gamma|t\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx|^\alpha\right) \\ & = & \exp\left(i\mu_m t-\gamma_m|t|^\alpha\right),\end{array}$$

where:

$$\left\{\begin{array}{rcl}\gamma_m & = & \gamma\sum_{n=1}^N|\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx|^\alpha \\ \mu_m & = & \left(\hat{Z}_{\eta,j,k}\right)_m+\mu\sum_{n=1}^N\psi^\eta_{j,k}(U_n)\int_{V(U_n)}dx.\end{array}\right.$$

We recognize the characteristic function of a symmetric alpha-stable random variable of parameters α, γ_m, μ_m, like in equation (4). □

H Proof of Theorem 5

Proof. Like in the proof of Theorem 2, we achieve a first-order Taylor expansion on Z:

$$\forall n\in\{1,...,N\},\ Z^\varepsilon(U_n)\overset{\varepsilon^\star_{1,n},...,\varepsilon^\star_{p,n}\to 0}{\sim}Z(U_n)-\sum_{m=1}^p D_mZ(U_n)\varepsilon^\star_{m,n}+(\varepsilon_{1,n},...,\varepsilon_{p,n}). \tag{47}$$

Using the equation (47), the wavelet coefficient $\hat{Z}^{\varepsilon}_{\eta,j,k}$ of Z^{ε} is:

$$\begin{array}{ccl} \hat{Z}^{\varepsilon}_{\eta,j,k} & = & \sum_{n=1}^{N} Z^{\varepsilon}(U_n)\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx \\ & \underset{\max_{n\in\{1,...,N\}}|\varepsilon^{\star}_{1,n}|+...+|\varepsilon^{\star}_{p,n}|\to 0}{\sim} & \sum_{n=1}^{N}\left[Z(U_n)-\sum_{m=1}^{p}D_mZ(U_n)\varepsilon^{\star}_{m,n}+(\varepsilon_{1,n},...,\varepsilon_{p,n})\right]\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx \\ & \underset{\max_{n\in\{1,...,N\}}|\varepsilon^{\star}_{1,n}|+...+|\varepsilon^{\star}_{p,n}|\to 0}{\sim} & \hat{Z}_{\eta,j,k}+\sum_{n=1}^{N}\left[-\sum_{m=1}^{p}D_mZ(U_n)\varepsilon^{\star}_{m,n}+(\varepsilon_{1,n},...,\varepsilon_{p,n})\right]\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx. \end{array}$$

Then, the limit of the random variable $\hat{Z}^{\varepsilon}_{\eta,j,k}$ when $\max_{n\in\{1,...,N\}}|\varepsilon^{\star}_{1,n}|+...+|\varepsilon^{\star}_{p,n}|\to 0$, denoted $\hat{Z}^{\varepsilon,\lim}_{\eta,j,k}$, is simply:

$$\hat{Z}^{\varepsilon,\lim}_{\eta,j,k}=\hat{Z}_{\eta,j,k}+\sum_{n=1}^{N}\left[-\sum_{l=1}^{p}D_lZ(U_n)\varepsilon^{\star}_{l,n}+(\varepsilon_{1,n},...,\varepsilon_{p,n})\right]\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx.$$

Let $t\in\mathbb{R}$. The characteristic function of the m-th coordinate of $\hat{Z}^{\varepsilon,\lim}_{\eta,j,k}$ is then:

$$\begin{array}{rcl} \mathbb{E}\left[\exp\left(it\left(\hat{Z}^{\varepsilon,\lim}_{\eta,j,k}\right)_m\right)\right] & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m\right)\prod_{n=1}^{N}\mathbb{E}\left[\exp\left(it\varepsilon_{m,n}\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx\right)\right] \\ & & \prod_{n=1}^{N}\mathbb{E}\left[\exp\left(-it\sum_{l=1}^{p}\left(D_lZ(U_n)\right)_m\varepsilon^{\star}_{l,n}\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx\right)\right] \\ & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m\right)\prod_{n=1}^{N}\mathbb{E}\left[\exp\left(it\varepsilon_{m,n}\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx\right)\right] \\ & & \prod_{n=1}^{N}\prod_{l=1}^{p}\mathbb{E}\left[\exp\left(-it\left(D_lZ(U_n)\right)_m\varepsilon^{\star}_{l,n}\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx\right)\right], \end{array} \tag{48}$$

since the variables $\varepsilon^{\star}_{1,1},\varepsilon^{\star}_{1,2},...,\varepsilon^{\star}_{1,N},...,\varepsilon^{\star}_{p,1},\varepsilon^{\star}_{p,2},...,\varepsilon^{\star}_{p,N},\varepsilon_{m,1},\varepsilon_{m,2},...,\varepsilon_{m,N}$ are independent. Then, using the equation (4) in the equation (48), we get:

$$\begin{array}{rcl} \mathbb{E}\left[\exp\left(it\left(\hat{Z}^{\varepsilon,\lim}_{\eta,j,k}\right)_m\right)\right] & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m\right)\prod_{n=1}^{N}\exp\left(i\mu t\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx-\gamma|t\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx|^{\alpha}\right) \\ & & \prod_{n=1}^{N}\prod_{l=1}^{p}\exp\left(-i\mu t\left(D_lZ(U_n)\right)_m\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx\right) \\ & & \prod_{n=1}^{N}\prod_{l=1}^{p}\exp\left(-\gamma|\left(D_lZ(U_n)\right)_m|^{\alpha}|t\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx|^{\alpha}\right) \\ & = & \exp\left(it\left(\hat{Z}_{\eta,j,k}\right)_m+\sum_{n=1}^{N}i\mu t\left(1-\sum_{l=1}^{p}\left(D_lZ(U_n)\right)_m\right)\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx\right) \\ & & \exp\left(-\sum_{n=1}^{N}\gamma\left(1+\sum_{l=1}^{p}\left|\left(D_lZ(U_n)\right)_m\right|^{\alpha}\right)|t\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx|^{\alpha}\right) \\ & = & \exp\left(i\mu_m t-\gamma_m|t|^{\alpha}\right), \end{array}$$

where:

$$\left\{\begin{array}{rcl} \gamma_m & = & \gamma\sum_{n=1}^{N}\left(1+\sum_{l=1}^{p}\left|\left(D_lZ(U_n)\right)_m\right|^{\alpha}\right)|\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx|^{\alpha} \\ \mu_m & = & \left(\hat{Z}_{\eta,j,k}\right)_m+\mu\sum_{n=1}^{N}\left(1-\sum_{l=1}^{p}\left(D_lZ(U_n)\right)_m\right)\psi^{\eta}_{j,k}(U_n)\int_{V(U_n)}dx. \end{array}\right.$$

We recognize the characteristic function of a symmetric alpha-stable random variable of parameters α, γ_m, μ_m, like in equation (4). □

I Chebyshev interpolation polynomials

Chebyshev polynomials of the first kind are described by:

$$T_n(x)=\left\{\begin{array}{ll} \frac{n}{2}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}(-1)^k\frac{(n-k-1)!}{k!(n-2k)!}(2x)^{n-2k} & \text{if } n\geq 1 \\ 1 & \text{else,} \end{array}\right.$$

where $x\in\mathbb{R}$. The roots of T_{n+1} are noted $x_0,...,x_n$ and they are defined by:

$$x_j=\cos\left(\frac{\left(j+\frac{1}{2}\right)\pi}{n+1}\right).$$

We want to interpolate the function z, knowing its value at some nodes which are the $x_0, ..., x_n$ translated and scaled. More precisely, the nodes are taken on an interval (a, b) wider than the support of z. We note them $\zeta(x_0), ..., \zeta(x_n)$, where $\zeta : x \mapsto \frac{1}{2}(a+b) + \frac{1}{2}(b-a)x$ is an affine function. Because of the discrete orthogonality of the Chebyshev polynomials, [21], one gets P_n, the Chebyshev interpolation polynomial of z:

$$P_n(x) = \sum_{k=0}^{n} c_k \delta_k T_k(\zeta^{-1}(x)),$$

for $x \in (a, b)$, with

$$\delta_k = \begin{cases} 1/2 & \text{if } k = 0 \\ 1 & \text{else,} \end{cases}$$

and

$$c_k = \frac{2}{n+1} \sum_{j=0}^{n} z(\zeta(x_j)) T_k(x_j).$$

Let

$$a_{k,j} = \begin{cases} (-1)^j 2^{k-2j-1} c_k k \frac{(k-j-1)!}{j!(k-2j)!} & \text{if } k \geq 1 \\ \frac{c_k}{2} & \text{else.} \end{cases}$$

Then

$$\begin{aligned} P_n(x) &= \sum_{k=0}^{n} \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} a_{k,j} \left(\zeta^{-1}(x)\right)^{k-2j} \\ &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{j=0}^{k} a_{2k,j} \left(\zeta^{-1}(x)\right)^{2k-2j} + \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \sum_{j=0}^{k} a_{2k+1,j} \left(\zeta^{-1}(x)\right)^{2k+1-2j} \\ &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{j=0}^{k} a_{2k,k-j} \left(\zeta^{-1}(x)\right)^{2j} + \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \sum_{j=0}^{k} a_{2k+1,k-j} \left(\zeta^{-1}(x)\right)^{2j+1} \\ &= \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{k=j}^{\lfloor \frac{n}{2} \rfloor} a_{2k,k-j} \left(\zeta^{-1}(x)\right)^{2j} + \sum_{j=0}^{\lfloor \frac{n-1}{2} \rfloor} \sum_{k=j}^{\lfloor \frac{n-1}{2} \rfloor} a_{2k+1,k-j} \left(\zeta^{-1}(x)\right)^{2j+1}. \end{aligned}$$

Therefore

$$\begin{aligned} P_n(x) &= \frac{c_0}{2} + \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{k=j,k>0}^{\lfloor \frac{n}{2} \rfloor} (-1)^{k-j} 2^{2j} c_{2k} k \frac{(k+j-1)!}{(k-j)!(2j)!} \left(\zeta^{-1}(x)\right)^{2j} \\ &+ \sum_{j=0}^{\lfloor \frac{n-1}{2} \rfloor} \sum_{k=j}^{\lfloor \frac{n-1}{2} \rfloor} (-1)^{k-j} 2^{2j} c_{2k+1} (2k+1) \frac{(k+j)!}{(k-j)!(2j+1)!} \left(\zeta^{-1}(x)\right)^{2j+1}, \end{aligned}$$

what may also be written with the power series form with a center $c = \frac{a+b}{2}$:

$$\begin{aligned} P_n(x) &= \frac{c_0}{2} + \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{k=j,k>0}^{\lfloor \frac{n}{2} \rfloor} (-1)^{k-j} \frac{2^{4j}}{(b-a)^{2j}} c_{2k} k \frac{(k+j-1)!}{(k-j)!(2j)!} (x-c)^{2j} \\ &+ \sum_{j=0}^{\lfloor \frac{n-1}{2} \rfloor} \sum_{k=j}^{\lfloor \frac{n-1}{2} \rfloor} (-1)^{k-j} \frac{2^{4j+1}}{(b-a)^{2j+1}} c_{2k+1} (2k+1) \frac{(k+j)!}{(k-j)!(2j+1)!} (x-c)^{2j+1}. \end{aligned}$$